```
   1 ATGAGCCAGC CCAGGCCCCG CTACGTGGTA GACAGAGCCG CATACTCCCT
  51 TACCCTCTTC GACGATGAGT TTGAGAAGAA GGACCGGACA TACCCAGTGG
 101 GAGAGAAACT TCGCAATGCC TTCAGATGTT CCTCAGCCAA GATCAAAGCT
 151 GTGGTGTTTG GGCTGCTGCC TGTGCTCTCC TGGCTCCCCA AGTACAAGAT
 201 TAAAGACTAC ATCATTCCTG ACCTGCTCGG TGGACTCAGC GGGGGATCCA
 251 TCCAGGTCCC ACAAGGCATG GCATTTGCTC TGCTGGCCAA CCTTCCTGCA
 301 GTCAATGGCC TCTACTCCTC CTTCTTCCCC CTCCTGACCT ACTTCTTCCT
 351 GGGGGGTGTT CACCAGATGG TGCCAGGTAC CTTTGCCGTT ATCAGCATCC
 401 TGGTGGGTAA CATCTGTCTG CAGCTGGCCC CAGAGTCGAA ATTCCAGGTC
 451 TTCAACAATG CCACCAATGA GAGCTATGTG GACACAGCAG CCATGGAGGC
 501 TGAGAGGCTG CACGTGTCAG CTACGCTAGC CTGCCTCACC GCCATCATCC
 551 AGATGGGTCT GGGCTTCATG CAGTTTGGCT TTGTGGCCAT CTACCTCTCC
 601 GAGTCCTTCA TCCGGGGCTT CATGACGGCC GCCGGCCTGC AGATCCTGAT
 651 TTCGGTGCTC AAGTACATCT TCGGACTGAC CATCCCCTCC TACACAGGCC
 701 CAGGGTCCAT CGTCTTTACC TTCATTGACA TTTGCAAAAA CCTCCCCCAC
 751 ACCAACATCG CCTCGCTCAT CTTCGCTCTC ATCAGCGGTG CCTTCCTGGT
 801 GCTGGTGAAG GAGCTCAATG CTCGCTACAT GCACAAGATT CGCTTCCCCA
 851 TCCCTACAGA GATGATTGTG GTGGTGGTGG CAACAGCTAT CTCCGGGGGC
 901 TGTAAGATGC CCAAAAAGTA TCACATGCAG ATCGTGGGAG AAATCCAACG
 951 CGGGTTCCCC ACCCCGGTGT CGCCTGTGGT CTCACAGTGG AAGGACATGA
1001 TAGGCACAGC CTTCTCCCTA GCCATCGTGA GCTACGTCAT CAACCTGGCT
1051 ATGGGCCGGA CCCTGGCCAA CAAGCACGGC TACGACGTGG ATTCGAACCA
1101 GGAGATGATC GCTCTCGGCT GCAGCAACTT CTTTGGCTCC TTCTTTAAAA
1151 TTCATGTCAT TTGCTGTGCG CTTTCTGTCA CTCTGGCTGT GGATGGAGCT
1201 GGAGGAAAAT CCCAGGTGGC CAGCCTGTGT GTGTCTCTGG TGGTGATGAT
1251 CACCATGCTG GTCCTGGGGA TCTATCTGTA TCCTCTCCCT AAGTCTGTGC
1301 TAGGAGCCCT GATCGCTGTC AATCTCAAGA ACTCCCTCAA GCAACTCACC
1351 GACCCCTACT ACCTGTGGAG GAAGAGCAAG CTGGACTGTT GCATCTGGGT
1401 AGTGAGCTTC CTCTCCTCCT TCTTCCTCAG CCTGCCCTAT GGTGTGGCAG
1451 TGGGTGTCGC CTTCTCCGTC CTGGTCGTGG TCTTCCAGAC TCAGTTTCGA
1501 AATGGCTATG CACTGGCCCA GGTCATGGAC ACTGACATTT ATGTGAATCC
1551 CAAGACCTAT AATAGGGCCC AGGATATCCA GGGGATTAAA ATCATCACGT
1601 ACTGCTCCCC TCTCTACTTT GCCAACTCAG AGATCTTCAG GCAAAAGGTC
1651 ATCGCCAAGA CTGTCTCCCT GCAGGAGCTG CAGCAGGACT TTGAGAATGC
1701 GCCCCCCACC GACCCCAACA ACAACCAGAC CCCGGCTAAC GGCACCAGCG
1751 TGTCCTATAT CACCTTCAGC CCTGACAGCT CCTCACCTGC CCAGAGTGAG
1801 CCACCAGCCT CCGCTGAGGC CCCCGGCGAG CCCAGTGACA TGCTGGCCAG
1851 CGTCCCACCC TTCGTCACCT TCCACACCCT CATCCTGGAC ATGAGTGGAG
1901 TCAGCTTCGT GGACTTGATG GGCATCAAGG CCCTGGCCAA GCTGAGCTCC
1951 ACCTATGGGA AGATCGGCGT GAAGGTCTTC TTGGTGAACA TCCATGCCCA
2001 GGTGTACAAT GACATTAGCC ATGGAGGCGT CTTTGAGGAT GGGAGTCTAG
2051 AATGCAAGCA CGTCTTTCCC AGCATACATG ACGCAGTCCT CTTTGCCCAG
2101 GCAAATGCTA GAGACGTGAC CCCAGGACAC AACTTCCAAG GGGCTCCAGG
2151 GGATGCTGAG CTCTCCTTGT ACGACTCAGA GGAGGACATT CGCAGCTACT
2201 GGGACTTAGA GCAGGAGATG TTCGGGAGCA TGTTTCACGC AGAGACCCTG
2251 ACCGCCCTGT GA (SEQ ID NO:1)
```

FEATURES:
Start Codon: 1
Stop Codon: 2260

FIGURE 1A

HOMOLOGOUS PROTEINS:

<u>Top BLAST Hits:</u>

			Score	E
gb|AAF81911.1|AF279265_1	(AF279265) putative anion transpor...		476	e-133
gb|AAF71715.1|AF230376_1	(AF230376) prestin [Meriones ungui...		471	e-131
ref|NP_000432.1|	pendrin [Homo sapiens] >gi|11421915|ref|XP...		451	e-125
ref|NP_035997.1|	Pendred syndrome homolog (human); Pendred'...		448	e-124
ref|NP_062087.1|	Pendred syndrome homolog (human) [Rattus n...		447	e-124
ref|NP_067328.1|	down-regulated in adenoma [Mus musculus] >...		434	e-120
ref|NP_000102.1|	down-regulated in adenoma protein [Homo sa...		418	e-115
sp|O70531|DTD_RAT	SULFATE TRANSPORTER (DIASTROPHIC DYSPLASI...		365	1e-99
ref|NP_000103.1|	sulfate anion transporter 1; Diastrophic d...		362	1e-98
ref|NP_031911.1|	diastrophic dysplasia [Mus musculus] >gi|2...		357	4e-97

<u>BLAST to dbEST:</u>

	Score	E
gi|8630793 /dataset=dbest /taxon=960...	523	e-146

EXPRESSION INFORMATION FOR MODULATORY USE:
library source:
<u>Expression information from BLAST dbEST hits:</u>
gi\|8630793 Human head-neck

<u>Expression information from PCR-based tissue screening panels:</u>
Human fetal lung

FIGURE 1B

```
  1 MSQPRPRYVV DRAAYSLTLF DDEFEKKDRT YPVGEKLRNA FRCSSAKIKA
 51 VVFGLLPVLS WLPKYKIKDY IIPDLLGGLS GGSIQVPQGM AFALLANLPA
101 VNGLYSSFFP LLTYFFLGGV HQMVPGTFAV ISILVGNICL QLAPESKFQV
151 FNNATNESYV DTAAMEAERL HVSATLACLT AIIQMGLGFM QFGFVAIYLS
201 ESFIRGFMTA AGLQILISVL KYIFGLTIPS YTGPGSIVFT FIDICKNLPH
251 TNIASLIFAL ISGAFLVLVK ELNARYMHKI RFPIPTEMIV VVVATAISGG
301 CKMPKKYHMQ IVGEIQRGFP TPVSPVVSQW KDMIGTAFSL AIVSYVINLA
351 MGRTLANKHG YDVDSNQEMI ALGCSNFFGS FFKIHVICCA LSVTLAVDGA
401 GGKSQVASLC VSLVVMITML VLGIYLYPLP KSVLGALIAV NLKNSLKQLT
451 DPYYLWRKSK LDCCIWVVSF LSSFFLSLPY GVAVGVAFSV LVVVFQTQFR
501 NGYALAQVMD TDIYVNPKTY NRAQDIQGIK IITYCSPLYF ANSEIFRQKV
551 IAKTVSLQEL QQDFENAPPT DPNNNQTPAN GTSVSYITFS PDSSSPAQSE
601 PPASAEAPGE PSDMLASVPP FVTFHTLILD MSGVSFVDLM GIKALAKLSS
651 TYGKIGVKVF LVNIHAQVYN DISHGGVFED GSLECKHVFP SIHDAVLFAQ
701 ANARDVTPGH NFQGAPGDAE LSLYDSEEDI RSYWDLEQEM FGSMFHAETL
751 TAL (SEQ ID NO:2)
```

FEATURES:
Functional domains and key regions:
[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 3
```
        1      153-156 NATN
        2      156-159 NESY
        3      580-583 NGTS
```

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 2
```
        1       45-47 SAK
        2      445-447 SLK
```

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 11
```
        1       18-21 TLFD
        2      158-161 SYVD
        3      240-243 TFID
        4      365-368 SNQE
        5      459-462 SKLD
        6      556-559 SLQE
        7      635-638 SFVD
        8      691-694 SIHD
        9      722-725 SLYD
       10      726-729 SEED
       11      732-735 SYWD
```

[4] PDOC00007 PS00007 TYR_PHOSPHO_SITE
Tyrosine kinase phosphorylation site

Number of matches: 2
```
        1        7-15 RYVVDRAAY
        2      447-454 KQLTDPYY
```

FIGURE 2A

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 10

1	77-82	GGLSGG
2	78-83	GLSGGS
3	89-94	GMAFAL
4	103-108	GLYSSF
5	335-340	GTAFSL
6	435-440	GALIAV
7	481-486	GVAVGV
8	485-490	GVAFSV
9	581-586	GTSVSY
10	681-686	GSLECK

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	51	71	0.893	Putative
2	82	102	1.020	Certain
3	107	127	1.729	Certain
4	130	150	1.497	Certain
5	186	206	1.723	Certain
6	228	248	1.517	Certain
7	256	276	1.898	Certain
8	288	308	1.252	Certain
9	338	358	1.568	Certain
10	383	403	1.304	Certain
11	412	432	2.345	Certain
12	469	489	1.997	Certain
13	619	639	1.146	Certain

FIGURE 2B

BLAST Alignment to Top Hit:
>gb|AAF81911.1|AF279265_1 (AF279265) putative anion transporter 1 [Homo
sapiens]
          Length = 738

 Score = 476 bits (1224), Expect = e-133
 Identities = 263/724 (36%), Positives = 428/724 (58%), Gaps = 36/724 (4%)
 Frame = +3

```
Query: 54    LFDDEFEEKDR--TYPVGEKLRNAFRCSSAKIKAVVFGLLPVLSWLPKYKIKDYIIPDLL 227
             L + EE R + P + R +CS A+ A++   LPVL WLP+Y ++D+++ DLL
Sbjct: 15    LNQEHLEELGRWGSAPRTHQWRTWLQCSRARAYALLLQHLPVLVWLPRYPVRDWLLGDLL 74

Query: 228   GGLSGGSIQVPQGMAFALLANLPAVNGLYSSFFPLLTYFFLGGVHQMVPGTFAVISILVG 407
             GLS +Q+PQG+A+ALLA LP V GLYSSF+P+ YF G   + GTFAV+S++VG
Sbjct: 75    SGLSVAIMQLPQGLAYALLAGLPPVFGLYSSFYPVFIYFLFGTSRHISVGTFAVMSVMVG 134

Query: 408   NICLQLAPESKFQVFNNATNESYVDTAAMEAERLHVSATLACLTAIIQMGLGFMQFGFVA 587
             ++  LAP+     A N+S ++ A +A R+ V++TL+ L + Q+GLG + FGFV
Sbjct: 135   SVTESLAPQ--------ALNDSMINETARDAARVQVASTLSVLVGLFQVGLGLIHFGFVV 186

Query: 588   IYLSESFIRGFMTAAGLQILISVLKYIFGLTIPSYTGPGSIVFTFIDICKNLPHTNIASL 767
             YLSE +RG+ TAA +Q+ +S LKY+FGL + S++GP S+++T +++C  LP + + ++
Sbjct: 187   TYLSEPLVRGYTTAAAVQVFVSQLKYVFGLHLSSHSGPLSLIYTVLEVCWKLPQSKVGTV 246

Query: 768   IFALISGAFLVLVKELNARYMHKIRFPIPTEMIVVVVATAISGGCKMPKKYHMQIVGEIQ 947
             + A ++G LV+VK LN +   ++ PIP E++ ++ AT IS G  + ++ + +VG I
Sbjct: 247   VTAAVAGVVLVVVKLLNDKLQQQLPMPIPGELLTLIGATGISYGMGLKHRFEVDVVGNIP 306

Query: 948   RGFPTPVSPVVLQWKDMIGTAFSLAIVSYVINLAMGRTLANKHGYDVDSNQEMIALGCSN 1127
             G  PV+P   + ++G+AF++A+V + I +++G+  A +HGY VDSNQE++ALG SN
Sbjct: 307   AGLVPPVAPNTQLFSKLVGSAFTIAVVGFAIAISLGKIFALRHGYRVDSNQELVALGLSN 366

Query: 1128  FFGSFFKIHVICCALSVTLAVDGAGGKSQVASLCVSLVVMITMLVLGIYLYPLPKSVLGA 1307
             G F+ + C++S +L + GG SQVA  SL +++ ++ LG  + LPK+VL A
Sbjct: 367   LIGGIFQCFPVSCSMSRSLVQESTGGNSQVAGAISSLFILLIIVKLGELFHDLPKAVLAA 426

Query: 1308  LIAVNLKNSLKQLTDPYYLWRKSKLDCCIWVVSFLSSFFLSLPYGVAVGVAFSVLVVVFQ 1487
             +I VNLK L+QL+D  LW+ ++ D IW+V+F ++  L+L G+ V V FS+L+VV +
Sbjct: 427   IIIVNLKGMLRQLSDMRSLWKANRADLLIWLVTFTATILLNLDLGLVVAVIFSLLLVVVR 486

Query: 1488  TQFRNGYALAQVMDTDIYVNPKTYNRAQDIQGIKIITYCSPLYFANSEIF---------- 1637
             TQ +   L QV DTDIY +    Y+ A++++G+K+   + +YFAN+E +
Sbjct: 487   TQMPHYSVLGQVPDTDIYRDVAEYSEAKEVRGVKVFRSSATVYFANAEFYSDALKQRCGV 546

Query: 1638  --------RQKVIAK--TVSLQELQQDFE-NAPPTDPNNNQTPAN-GTSVSYI------- 1760
                     ++K++ K  + L++LQ++ +     P     N TS+ +
Sbjct: 547   DVDFLISQKKKLLKKQEQLKLKQLQKEEKLRKQAASPKGASVSINVNTSLEDMRSNNVED 606

Query: 1761  -----TFSPDSSSPAQSEPPASAEAPGEPSDMLASVPPFVTFHTLILDMSGVSFVDLMGI 1925
                  S D  A +     ++AP + S + A   P    FH+LILD+ +SFVD + +
Sbjct: 607   CKMMQVSSGDKMEDATANGQEDSKAP-DGSTLKALGLPQPDFHSLILDLGALSFVDTVCL 665

Query: 1926  KALAKLSSTYGKIGVKVFLVNIHAQVYNDISHGGVFEDGSLECKHVFPSIHDAVLFAQAN 2105
             K+L + + +I V+V++ H+ V + + G F D S+ KH+F S+HDAV FA +
Sbjct: 666   KSLKNIFHDFREIEVEVYMAACHSPVVSQLEAGHFF-DASITKKHLFASVHDAVTFALQH 724

Query: 2106  ARDV 2117
             R V
Sbjct: 725   PRPV 728 (SEQ ID NO :4)
```

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00916	Sulfate transporter family	254.5	1.5e-72	1
PF00189	Ribosomal protein S3, C-terminal domain.	3.3	8	1

FIGURE 2C

Parsed for domains:

Model	Domain	seq-f	seq-t		hmm-f	hmm-t		score	E-value
PF00916	1/1	187	497	..	1	328	[]	254.5	1.5e-72
PF00189	1/1	651	661	..	79	89	.]	3.3	8

FIGURE 2D

```
   1 CTGGGTTCCT ATGTGGGGAG GTCATGCTCC CCACTCATTG AGCCCCCCCA
  51 GGCAAACCAC CTGGACAGCC AGACCCATGC AGACTCTGGA GCAGGTGGAG
 101 AGGAAGAGTG AGACCACCCC GCCTCACGGG CGGTGAAGGG CCGGCAGCCT
 151 CTGAATAGTC TCTGCTAGGA GGTAGAAAGC ACCCTCCCAT CTTAATCATA
 201 GTAATCATCG CCACTACCAT TTACTGGGTG CCTATAAAAG GCCAGCCTCT
 251 TCATACACAT GATCTCACTG AATCCTCATA GCATCTGCCT GCGACTGTTA
 301 TTATCCCCAT TTACAGATGA AGAAACTGAA TCTTTGAACC CAGGTCATCT
 351 GGCTCTCAAA CTTGTGCTGT TTTCCCTAAG CCACCCGGTC TCTCATTTCT
 401 CCCACTGAAA TGTCTCACAT GCCATTGCCC TTACTCATTT CTGCCCATGT
 451 CTCCTCCAAA ACACCATTTA TCAATTCGCT CAACAAGTAT GTGTTGAGTA
 501 CACACTAAGG GCCAGGCGAG GGGCTGGGCA CAGGCGCTGG GGGTAGGTTC
 551 ATTCTCCCAC CTTCGCTTCT GCTGGGTATC ACCTGTGGGG TCTTGCCGGG
 601 CATCCCACCC TCACCTGTAG TTCAAGTGGA CCTTGGGATC CCAAGACCAA
 651 ATGAATGGAA TGCACCAGCC CAGCCTTCAC CAACTTGAGC ACAATCTTAT
 701 TCATAATAGA AACTCACATT TGCATCACAC TTTACATTTT ACACAACCCC
 751 TTCTTATCCA TTAACTCATT TGATCTTCAC AACAACCCTG TGAGATATGT
 801 CTGTTACTCC CACTTTAGTG ATACAGAATC TGAGGTTTGA AAAGTAATGC
 851 TGACCATTCT GCCTCATTAA TAAAAGCAGG ATTAACCCAG GCTCCTGGAC
 901 CCTTCCACAA AAGGCATTAA GCAACCTGCT CCCCTCTGAC AACCTCCCCT
 951 GTCACCCAGG CTCTCCTCTG GGAAGTTGGG GGCATCTCTA GCCCCCAAGT
1001 AGTTACTCAT TTTCAACCCC ATCTCAAATC TTTTGCCAAA CTGGCCACAG
1051 CCACCCCACA CTCCCCACCT CCCAGATACA AATCCTCACT CTAAGCCTTC
1101 CCCATCTCTT TCTTCTCTGT CCTTCTTTCT CTGTGGTCCT CTGAGCAACT
1151 TCTCCCAGCT CTGGGAGGTA GAGGGGAGGT GGGAGACCCA GTAATTGGAA
1201 GAGGGAGGGG GAAAGGTTCC TACAGGGAAC TCCTCCGGGC CTCAGGGGCC
1251 CTGGCACTCA GCTCTGCCCA TCTCAGCTCC TGGAACGTCA GCCAGGTTGC
1301 GCAAAAAGTG AGGAGGAGAG GAGCGGCAGT ACACAAGGGT GGGGGAAAGA
1351 TTAGGCACAG GAAGCCGTGG GAGAGAGAGC CGGCAGGTGG ACCATCCTGG
1401 TTTCCCCACA CACACCATTG TCCCCCTGGG AAACCTGTTG GTGAAGTTCT
1451 AGATGTCTTA TCCAAGAAGG GTCCTCTTGA GGTCATCTCA GCTATCCCCC
1501 TGCCTCTAGG CAAGCTGTTT TCTGTTTCTT CCAAGCTGAC TGGCTGAATG
1551 GTAGGAGCCT TTCTGCCAGG GAAACTAAGG TCTGGGAAGG GAGTATGGCT
1601 TGTGGGGACA CCAGGGGTCA GGGGAGGGGA GGGTCCACCT GCTGAATCAA
1651 GTGGGGCCTC CTGCCCTCGT GATTCCCCTT TGCCTGGTGC TCAGTGGGGG
1701 TGATGGTGAC GCCACAGGTG TGGAGTGCCA GCCACGTGCT GAGCGCCAAG
1751 CAAAACAGCC AGGGTGAGTC TATGCATCAT CAGTGCCTGG GAAGGAAGGC
1801 CACTGCGAGC AGGGAGTCTG ACGGAAAAAC TTGACAGAGG GAAGGGAGGC
1851 ACCTTGCTTT ATCGGGGCGG GGAAGGCCAA AATAAAACTC TGCTACTGCA
1901 AGGACCAGAG AGAGAAGGCC TGGGCTGGCA CTAGGGAGGG ATGTTCCCTC
1951 ACCCTCCCCT CCTCTGCTTC TCCCAAAGCT TGTAAATGCC CCAGATATGA
2001 GCCAGCCCAG GCCCCGCTAC GTGGTAGACA GAGCCGCATA CTCCCTTACC
2051 CTCTTCGACG ATGAGTTTGA GAAGAAGGAC CGGACATACC CAGTGGGAGA
2101 GAAACTTCGC AATGCCTTCA GGTAACTGGT CCAGAGCCCA GACTTCTGCC
2151 TCCTCTGCTC CCTACCAAAA TCCTTTCTGC ACCAGGACAC GGCTTCTGCA
2201 CTGGTATCCC TAAGATGGGG TTAAGGGAAG CCCTGGGGAA GTGAGGTTCT
2251 GAATGATGAA TTTAAGATCC TACAACCTCA TCTGTACTGA GACCCCCAGG
2301 GAGGATGGGG AGCAGGAGCA AGAACCATCC AGAAGGGTTA TATGGCATTC
2351 CCAAACCCCT GCATGGCATC TCCCATATTC TCAATTCACC CGGGTCTCTC
2401 TGGGTTTGTT AAGGCATGGT AGATGAGCAT CTACGTTATG GAGGGGTGGG
2451 GAGCATCAGA GCCCTTACTC CATGCCCTGT TCCCTCCTTA CAAAAAATAC
2501 CTGAAGTTAC CATCACCCCA GGTTCTTTGT CCTTTCCCTC CCGGATGTTC
2551 CTTCCTCCAC TTGGTCCAGA GAATGCCAAA AGGAGGCCCT AAATTTCTGA
2601 ACTTTCCTGA GGGGACCTAC CAGGGTGTAG TCCTACCAGC GCCCAGGGTC
2651 TTTCCACTCT CATCTCCCTG GAAATGCGAT GGTGGGTATG AAACCTTGTC
2701 CCTAAGTAGG CGCTACACAA GGTGATCCAT ACCCACACCC CAGGAGGCTG
2751 GGGCTGCGGG TGTCACCCTC CCCATTCCCA GACTCCTGGC AGACCTCCTC
2801 TGGCCCAGCT ATAGGCCAAC TCACTCTCCC TCACTCCCTT GGGGAAACGG
2851 CTGATTCAGT TACCTGGATT GAGGTCACTG GCAATGGCTG AAGTGGAGAC
2901 GCAGGTGGAA CTGGTTCAGG CCGGGGGAAT CACCCACTTG AGTTTGTACT
2951 AAAAGCCCCA GCCCAGCCCT GTTTCTCTTG GGAGGCTCCA TTTCTGCCCA
3001 GTTACAGTCT GTCCTCACAG CTGTGCTCCT CAGACAGGTG GTCTCTGCCA
3051 GTCTTTGTGC CCAAGACTTT AGGGCACAAA GTCTGAGGAT GAGAAGATCT
3101 GCTATTGTCC TAAAAGATTA GGATAATGAA AGCTGTAAAG GGATATAGCA
3151 AACTAACAAT TCCTATGATA CTGGCATGAG AGCCTTGAAC AGTGCCTGGC
3201 ATAGAGAAGG TGCACCAATA AATATTTGTT TCATGAATGA ATGAATGAAT
3251 GAATGTCTAG AAAGCTAATC CCTCTCAGCC TCTGTTTCCA GTTCTTCTTT
3301 CAAGCTTCAG ATTGCTTTGC CCAACATACA GCAGACTTGC AAGTAAGGTT
3351 GGGCATGGAC TAGCCCTCAA ATGAGTTGTT TTTCTTTCCC TAGCCAGCTC
```

FIGURE 3A

```
3401 TCTATTCATA AGTCCGGCTT TCTCTGCCAC AAACAGACCT GATGGAGCCC
3451 CTGCAGGGCT GGTTCTCTCT TCAAGCAAGG CTTTAGAGTT GCATTAAGCA
3501 ATTTATCCCC CGTCCACCTC CCCTTCCAGC ATCCCAGGGA TGGCAGAGGC
3551 ACCCATGAGC CCCAGAAGGG ACAGGGGGTA AGATATTGAT GATGATGCTT
3601 TTTCTTGGAG TGTTAGTTGG AAGAGAAAAT CTGCCCAGAC TTTCCAAGGT
3651 ACAAAGCATT GTCTTTGTTG GTTTCAGTCT TGGGTGACAT CCAGGGGACC
3701 GAGTGTCAGG GAAACTATTG TTGAGCAAGA GCAAAGAGCA GGAATTGGTG
3751 CTGGGCAGGA AAGGAAGCCT CATCAGAGCA GGCCAGTGAG TCACCAAATG
3801 GGCCCTAAGT ATTTGAGTTC CCTCAACTGG GAGAAGGAAA GCAAATGCCC
3851 CTCACCCACT TCCAGTCATC AATCCACCGG CTGTCACCCT TGAGTTTGTA
3901 AGCCCTTGTT CCTACCGCTC CTGAGTTTCT ATGAAAGGAC CTTGAGGTGT
3951 TCAACAAACA GGGAAGGGAT CAACTCTCCC CACCCTGCGT TGACCAATGA
4001 ATTCTTCCCT CCTCTGCTGC CCAGTGAATT AACAGGAGAA AGAACTCCGG
4051 TATTGGAGTT ACCACACATA AAGGATAGTG AGTCAGCAGA GTGCACCCTG
4101 CAGGAACAAT AGAGCCTTCC TTTTCAAGGA AGTTCTAAGA AAAATGGCAG
4151 CAGGCAGGCC CCACTCGGGT GTATTCACTC ATTCATTTAT TCAACAAATA
4201 TTTACTAAGT GCCCCTGTGC AAGGCTCGAG GTGTACAAAG ATGAACAGGA
4251 GAGCTAGACT TCTTGCCATG CGTGGTGGGG TTTGCTGCCT AGTGGGAGAG
4301 ACAGACAAAA AGCAAGGAAT GCACACACAG GATGCACACA CAGCGGCAGG
4351 AACCAAGGTG CAGTTACCCA GGCCTGGGAT CAGACAGACA GGACTCAGAG
4401 GAGACTTTCC CAGAGAAAAG CCATCTGAGC CAAGGGATGG ATCTGATACC
4451 TCCGAAGGCT GAGCCACCAT AACACTCATA CCTTTAAGCC AAGTCTTATA
4501 AACTCCCCAG GTAAGCAGCT GGCAGTCAGA AGACCTCCAG CTAATGCCCA
4551 GGACAAGTTG ATGAGCTCTC AAGAAAAAGT TCCTGCCTTT TCTTCTCAAT
4601 ATCCCTGGCA CACAGTTCAG TGAATTTTGA ATGAACCAAT GAATGAAATG
4651 AGCAGGATAT GATAATCCCT CTCCAACACG GAATGTCCAA GCCATGCAGA
4701 GCCGACTGGA AATTTTCCCC GTTCCCTTCC AGATGTTCCT CAGCCAAGAT
4751 CAAAGCTGTG GTGTTTGGGC TGCTGCCTGT GCTCTCCTGG CTCCCCAAGT
4801 ACAAGATTAA AGACTACATC ATTCCTGACC TGCTCGGTGG ACTCAGCGGG
4851 GGATCCATCC AGGTCCCACA AGGTGAAGGG GCTCCTTCAG CCAGGCCTGG
4901 ATTGCCACTC CCCTCACCAT TCCTCTCCTC ATCCCCACTC CATCCCTCTG
4951 TGATCCCCAT AAGCTAGTCA TGCTGCTGAG CTTCAGTCTC GTTGTCCTCT
5001 GCAGGCATGG CATTTGCTCT GCTGGCCAAC CTTCCTGCAG TCAATGGCCT
5051 CTACTCCTCC TTCTTCCCCC TCCTGACCTA CTTCTTCCTG GGGGGTGTTC
5101 ACCAGATGGT GCCAGGTAAG GCCTCTCCCC TCTGGGCAGG CAGGATGACC
5151 CAGACCACAA GGATGGGAGG TGTGGCAAAG GGGCCTCGGG AGATTTTCCA
5201 TCTGCATTCT CCTGGAGTTG TTCCTGGTCA GTCCTAGGGG AATGGTCACT
5251 GTGAATGTCA TTTCCAGGTC CTCGGTGACC TTGGAGAAAC CACTGAGCCT
5301 CTTTGAGTTC AGTTAGCATT ACCTGTTCCA TCTTCCTCCT AGGAATGAGA
5351 GGAAGACTTA GCAGAACAAG ATATACCATA TGCTATAACA TGCTTAAACA
5401 GATGTGAGAA ATCACCATCT AACTCCCTGG TTGGTCCCAG CCGGCCACTA
5451 CAGGGACATT TGGACTTCTC TGGTGCTAAG TGAGATGGAG GAAAGCCTGG
5501 TCACAAGGGC TGGTTTCTGG TTCAGGCTCT GCTTATATTT CTTATTTCTG
5551 AGTTCATTTT CTCACGTGTC CTGTATGACA ATATTGACCA TTGGGGTAAA
5601 AGCACCTTGA AAAGCATAGA TCATGGTTAG AGTGAGTGGT TGTTATTATT
5651 GTGTTGGAGA AGAGCCTTGG AGGTGCAGGG ATCCATCCCC CTGGGGTCGG
5701 GAAGCATTCC TGGGCCCCTT TCTGGTTTCC ATCGGTGTGG TTCAAACCTC
5751 TGATTTTTGC TGGCTGGGTG GGGCACCACA GGTACCTTTG CCGTTATCAG
5801 CATCCTGGTG GGTAACATCT GTCTGCACGT GGCCCACAAG TCGAAATTCC
5851 AGGTCTTCAA CAATGCCACC AATGAGAGCT ATGTGGACAC AGCAGCCATG
5901 GAGGCTGAGA GGCTGCACGT GTCAGCTACG CTAGCCTGCC TCACTGCCAT
5951 CATCCAGGTG AGGGGGCAGC CCCCAACCCT GCTAGAAGGG CATCAGACCA
6001 CCCTGCCCCT CCCTCAAAGC CTTAGCTTTG ATGCTAAATC TGATTTAGGG
6051 GGCTGGGTGT GGAGGCTCAT GCCTGTAATC CCAGCACTTT GGGAGGCTGA
6101 GGAGGGTGGA TCACTTGAGG TCAGGAGTTT GAGACCACCT TGACCAACGT
6151 GATGAAACCC CATCTCTACC AAAAATACAA AAATAATCCA GGCTTGGTAG
6201 TATGCGCCTG TAGTCCCACC TACTCAGGAG GCTGAGGCAG GAGAATCACT
6251 TGAATCCGGG AGGCAGAGGT TGCAGTGAGC TGAGATCGCG CCACTGCACT
6301 CCAGCCTGGG TGACAGAGCG AGACTCCGTC TCAAAAAAAA AAAAAAAAAA
6351 AAAAAAAAAA CCCAAGTTAG GGCTCACCTC CTCCCTCCTC CCCATCCCAG
6401 GGCTAAAGTG AACCTTGAAA ATTAACAGTA TCTCCTCATC TGCATGTAGC
6451 AGGACCATAC AAAAAAACAA CAGCTGTACC TGGTTAAACT GTCCTGAGCT
6501 TTAAACCTGT AAAAGACTCA CAGCCTCTCT CCATTATCCC GTGGAGAAAC
6551 CCAACTCTCT GCCAGCATAG TCTTGCAGAC TGCTAATTTT CTCTAACATC
6601 CCTCACTCCG CTCCAGCCTC CTCTGCTCCA AGCCACAGCA GCAGTTGCAC
6651 AACATAAATT GAGCTTCTGC AAATGGTTGC AAAGGATTCT GCTAGGTTTT
6701 ATGAAGGGAA GCACAACATG ACAGAATGCA AGAGCAAAAC ACAGTCCCAG
6751 AGAGCGCCTT TTCATTCACT CATTCATTCG GTTTTGTGCC AAGAACTAGG
```

FIGURE 3B

```
 6801 CTAAACCCTG GGATACAAAG ATAAGTAAGA AAGAGGTCCA ATTCACAAGT
 6851 TGCTCACAGC CCAGCAGAGG AAGGAGCCAT GTCAACAGAT AAATTTGTAT
 6901 GCAGTGAGAT AAGCAGCAAA GTAGAGCCAT GTACAAAGAC TGTAGGGACA
 6951 CAGAGCAGAG TCACGGAGGA CCTCAAAGAG GAGGTGACAC TCCACCTCTC
 7001 TTAAAGGATG AGAACTTAAC CAGGAACAAG GTATACAGAG GATGGTCCAG
 7051 GCAGAAGGGA ACAGTGCCTA AAAACACTGA GGCCTGAGAG AGTGTGATCT
 7101 GCGCAGGCAA AGTAAGGGGC TTGGTGTGGC TGGAGGGTAG AGGGCCCAGA
 7151 AGAGGATGGA AAAGTAGGCA GGAGCCAGAC AATGAGATCT GGGGTCTGTT
 7201 CTCTGACAGC GACTTTGGGT CTGATTGGCA GTTTATAAGG ATCGTTTGGG
 7251 CTACACAATG ATGAGTGGGA GGTGGATTAG AATCAAGGCA GGGGACCTGT
 7301 TGGGAGACTC TGCAGAGGCC CAGGCAGGAA TAATGCAGGC GAAGACCAGG
 7351 TAGAGAAAGA GATGGGGCTG GACTTGAAAA GAATGTTTTA CCAGGAGCTT
 7401 GGTGATAGAC TGGATGTGGG AGGTAAGGGA GGATGACTCT CAAGTTTTTG
 7451 GTTGGGCAAC CAGGTTAATG ATGGTGTCAT TTACTGAGAG AGAAAACACT
 7501 GGGGGAGGAC TAGACTTATT TTACAGATAA GCCAAAGCCA GAGAGGTGAT
 7551 GTGACAGAAA GGCCCATGCT CTAAAGGAGC TGAAGGTCTG ATGGCAGCCA
 7601 TGTAGAGCAC AGTGAAGGGC AGGTGAAGGT CACAGATGGT CCAATTCCCT
 7651 CAAGCTACTG CTACGCTAGG ACTGCACGGA GCTCCAGACC TGCGTGTGTG
 7701 TGGGGCGGGT CGTTGGAACT GCTGAACCAC ATTGGTCTTC CGCCACCAAC
 7751 CACCCTTTTC CTCCTCTCAG ATGGGTCTGG GCTTCATGCA GTTTGGCTTT
 7801 GTGGCCATCT ACCTCTCCGA GTCCTTCATC CGGGGCTTCA TGACGGCCGC
 7851 CGGCCTGCAG ATCCTGATTT CGGTGCTCAA GTACATCTTC GGACTGGACCA
 7901 TCCCCTCCTA CACAGGCCCA GGGTCCATCG TCTTTGTGAG TCTGGGGATG
 7951 CACCCCTGCC ATTGGAGCAA GGCTCCAGCA GACACATGAG GAGGATGTAC
 8001 TGTTTTAAGA TGTCGTGAGC TCCTCATTGC AAGGGCTGGC TTAGCTGTTG
 8051 TTCAGAGAGG ATTCTGAGGG GGTTTCTGTC TTGGGAGGGT CAAAGTCATG
 8101 ACTCACAGAG GTTCTTGGTA GTTAATACCT GCAGAAAAGA GCTGTACATT
 8151 CTCCGCCAGT TCCCCATTCT AGTGCCTCAA CCCCTCCCTG CCTGGAAAGT
 8201 CCTGCCTTAT GTCTAATCTC CATCCCTCCT CCTTCAGCCC AAACTCTTCT
 8251 AAAGAAAAAG AAAGCATTCC TTTTCTAGCA CAAGTTCCCC ATGTGCCTTT
 8301 TGGGAAAGGG CGGTGGGCGA CGGGACAGGG TTCCTGATCA GGGTTTTAAT
 8351 TCTGTCTTGG TGTGCCTCCA TTAGCTTTGA TGGCATCCCT TCCCTGGGTC
 8401 AGACACCCAA AGGTGGGGTA TTATGGGAAG AAGGGGTGGG AGCCTGTGAG
 8451 CATGATGCTC TTTCCCCCAG ACCTTCATTG ACATTTGCAA AAACCTCCCC
 8501 CACACCAACA TCGCCTCGCT CATCTTCGCT CTCATCAGCG GTGCCTTCCT
 8551 GGTGCTGGTG AAGGAGCTCA ATGCTCGCTA CATGCACAAG ATTCGCTTCC
 8601 CCATCCCTAC AGAGATGATT GTGGTAAGGA CCTTGTTCAG AGCTGGGATG
 8651 TTGGGGGGCC AGGCTGTGAG ACGAGGAAGC CCCTACCTTT CCTCACCCCA
 8701 TCCCCTCAAC TGGCAGCCAG TGGGACAGGA AGTCAGTTGT GAATCCATCC
 8751 CATCCCCCGT ATGTGGCGTT TCCTCTCTTT CTACTGCTCT AATAATTCCC
 8801 CCTAAGGAGG CAGGGGAGTG GGATTCAGGG TCCCCAGAGA AAAGGGAGAC
 8851 TTGAGAGAGA CGCCTGCCCT GGCCCCACCT TAGGGCCAAT CCCCATTCTC
 8901 CACTCTGGGG TTTGCAGGTG GTGGTGGCAA CAGCTATCTC CGGGGGGCTGT
 8951 AAGATGCCCA AAAAGTATCA CATGCAGATC GTGGGAGAAA TCCAACGCGG
 9001 GTGAGTCCAG GTGGCCCAGA AGCCTGGCCC ACCCGCACCT CATGCCCCAC
 9051 TAAGGCCTGA GCTCGGAGAG GGAGACAAGA TGAACTCTAT GAAAGTGCAG
 9101 TCGAAACTGT ATGACACTGA CCATGTATGA ATTATTACTA TTACCGTTTC
 9151 CTGAGAAGGG CCGCACAACC AGCCAATGTA GGCTATTTTA TGAGAAATGA
 9201 GTCTTAACTG CCACACTCCC CTTATAAATC TCATTCAACT GATGCTGTTA
 9251 AACAAAGCCT CTCTGAACAG CCGCTTGCTG GCTCTTTGCC TTGCTCTAAT
 9301 GCATTGGTTC TTTGTCCATG TAGAAAGGGA ACTATTAGGT TCAACCAGAT
 9351 TCATGAAGCA TCCACTCTGT GCCAGGCACC ATGCTGGGCC CTGGGAGGAG
 9401 AGGGGTGACG CTTGTCCTGC AGGGTTGGAA CAGGCAAGGG AGGGAAGACC
 9451 ACATAGCACC AAAGGTCTAG GGGTCTGTGG ACTCGTGAGC ATACAGGGTT
 9501 CAGAATCTGG GAGTTAACAA ACGAGGCCCT ACCACATACT GGCCCGGGGA
 9551 CCTTGGGCAA GTTAGGTTCT CTCAGCCTCA GTTTCCTCCT TTGTAAAACA
 9601 GGAGTGATGG TCCCTACCCT ATGGGGTGGT GCTGAGGATT CAGACTGGAT
 9651 GGGATAACTT AGGCAAAGAT CCCGGCACAC CATGGGGGCC TGGCTGGTCC
 9701 CTGTGGGCTG GTGAAGGACT TGGCTGCCCT CCCCACTCAC ACCCTTGGGT
 9751 TCTGCCTCCT TCCTGGCTCC TCGGCAGGTT CCCCACCCCG GTGTCGCCTG
 9801 TGGTCTCACA GTGGAAGGAC ATGATAGGCA CAGCCTTCTC CCTAGCCATC
 9851 GTGAGCTACG TCATCAACCT GGCTATGGGC CGGACCTGG CCAACAAGCA
 9901 CGGCTACGAC GTGGATTCGA ACCAGGTAGC TCTGGCCACC CCCGGCAGGA
 9951 CTGGGCAGGA CAGGTCAACT CAGGCCTGGC ATGACATATC TTGGGTGGGG
10001 AGATCATTGG GCTGAGGTGA GGCAGGCTGC CTCGAGTGTG GGGGATAGGG
10051 GGTCCTCTGA CCCTAAGAGG CTGACCTCCT CTTGACTGGG AATGTGTGAC
10101 TTTATAGCCA CTGGGTCACT CTCAGGTCTT AGGCCCACAG TCCAGCTTGC
10151 ATGCCTGACT GCACTTGGTC CCCGTGCCCC CCAGCCCCAC ACTGGCTTCT
```

FIGURE 3C

```
10201 AATCCTGTCC CCTCCCTGCA GGAGATGATC GCTCTCGGCT GCAGCAACTT
10251 CTTTGGCTCC TTCTTTAAAA TTCATGTCAT TTGCTGTGCG CTTTCTGTCA
10301 CTCTGGCTGT GGATGGAGCT GGAGGAAAAT CCCAGGTGAG CCTTGTTCTA
10351 GGGGAGTTGG GGGGAGGTGG TAAGAGAACA GTTGCCCCAA AAAAGCCTGG
10401 GCACTGCAAG CCAGGCCAGC TCTTCTCCGA CCCCTTCTTC CCGTACTTAG
10451 TCTCCACTCC ACCAAAGCCA TGGATTGGAA ATAAATCAAG AGCAAAAATT
10501 TCACACCTTC CCTCTATCCC CAACTCTTTC TCGGAATAGG TGGCCAGCCT
10551 GTGTGTGTCT CTGGTGGTGA TGATCACCAT GCTGGTCCTG GGGATCTATC
10601 TGTATCCTCT CCCTAAGGTA AGAGCCCAGC CATCGAGCAG AAGTCAACGA
10651 AAGACTCCAA TAAGAACAAT CCCTGAGAGT TGTGTGGCAC TTTACGGACC
10701 ACAAAGTGCC ACTGTTGTCA TACTTAGTCT CAACCACAAA CTGTGAGGTA
10751 GACAATGCAG GTTTTATCCT CCCCATTTTA CAGGTGAAGG AAACTGAGTC
10801 TGAGAGTCTA AGTAACCTTG TCCATAGTGA GGCAGCTTAC AGCGCAGGGC
10851 TGGTCCCAAA CTCCAGCCTT CTGGCCTCAG AGTCTAATCC CTAGGCAACA
10901 TTTGCACCTA CCCACGAGTA CCAGGCTCTT ATATAGCCCA GCTAGGAGGG
10951 CTCTAGGCAT GCGTCATTTA GAGATGAGGG AAGAGAGATA GGGAAAGGAT
11001 GGGGCCAGGA AGGACCCCAT GGCTCTAACG CCAGCACTTT CCAAACCTAA
11051 GGTCGAATGC AGAGATTTGG GGGATCAGCC AGGGGAGGTG TTCCAGAACT
11101 CCGTCTCTGT CCTGCCAGGC CTTGGGGTCG GGTATGCGCA GGAGGGCAAA
11151 AAGAAGGGGA GACCCTGGGG TCCTGGAGCA ATGTTCTGCT TCTCTAGTCT
11201 GTGCTAGGAG CCCTGATCGC TGTCAATCTC AAGAACTCCC TCAAGCAACT
11251 CACCGACCCC TACTACCTGT GGAGGAAGAG CAAGCTGGAC TGTGTAAGTA
11301 TCGGGCAGCC TCTGGGTACT GGCCATGCCC CTGCCCTCTC CTCCAACCCC
11351 ACAGCCCTGT CAGCCCTGTC CTAACAATGA ACCCTCTAGT CTGCTGCTTC
11401 CTAATTAGCA TGAGATGAGT GGTTAAAAGT CCGAGTTTCG AAGTGAAACA
11451 TCCTATGTTC AAACCCTAAC TCAGCCATCT GCTGGCTCCA TGGCCAATAG
11501 CAAGCCCCTT AACCTTTCCC AGTCTTGGTG TCTTAACTGG GCAAATGGTT
11551 ATTTTATGCT CTCTGCCTCC CAGGGTTTTC TATGAAGAAG AAGCAAGGTA
11601 ATACAAGTAA ACATGTTGTC TACATCGTAT TTTATACTCA ATAAAGCTTA
11651 GCTATGACTA CTTTATGACA TACAGCTTTA AAAAACAAAA GGAAATAGTT
11701 TGTATTTTAA AAAAAAACCT AGAACATAAA GCCAGAGGAC CAAAATCTTG
11751 AGCAAGTTAC TAGACTTCCC TGGGGTTCTA TTTCCTCATC TGTAAATGGG
11801 GGTGAGACTC ATGCAGTCAT GGTTGCGTCA AACGCTGGTT CCGAGGATTA
11851 AATGAGATCC CAGTGGGAAA ACACCGCATG AGCGCAAACA TTCTGCAAAC
11901 ATGACTTATT GTCCTGATTA GTCACACACT CCACCGCATC ATCCGCTGGG
11951 CATAGTAATG AAGGCCAGTG TGTTTTGACG ACACTGCCTT CTCTCCATTT
12001 AAGCCCCACC ATAACCTATG GGAGAGGATT TACTAAACTT TCTTAACGGT
12051 GATGAAACCA AGGCTCAGAA TGGTTAAGTA AATTGTCAAA GGCCACAGAG
12101 GTAGGGAGTG GTAGAGTCTG GATTAAAACT CCAAGTCCTG GACTCCAGAC
12151 CTCTAGGCTG TACTGTCTCA TAGGGAAGGC AGTCTCACCC ACCTAGGGCA
12201 GAGAAGAAAA TCCTTAAAGC CAGAGAAGTG AGTGGCTCAT CTGTGGTCAC
12251 CCAGAGAGAC AGTGATGAGG ACAGGGAGAA AAATTATACC TCAGTTCCCA
12301 GCCCAAGGAT CTGCTTTGAC CATAACCCAA CAAGCCCCCG CTATGGTGGT
12351 ATTTCCTTAG GTTCATATGG CGGCTTTTGT TTCCATTTGA TCTTCACAGC
12401 AATTCTCTAC AGGAATCTGG GCAGATTTAT TTCCTTTAGA GGAATTTCCA
12451 GGTCTTAAAA TCTATAGGGG GCAACTATCA AAACTTCACC CAATGTTGCC
12501 CCCTACCCAC ACACAAAACC AGGCCCCCAG CCGATCAGAA AGCACTGCTG
12551 AGCTCCTGTC AGGGCCCACG CAGCTCGCTG TGAGACAGAG AGAGGGAACT
12601 CACATTTATT GATCACCTAC TGAGCATCCA TCACTAGGCT AGGACCGTCA
12651 CATTCCTTAA CTTTTGAATC CTTTCATGAG GTAGGCATTA TTATTCTCCT
12701 TTTGTTTCAC ATAGCCATTA AAGAACAAAA TTTGGGGCTG GGTGTGCTGA
12751 CTCACACCTG TGATCTAGCA CTTTAGGGGG CTGAGGCAGG AGGATCGCTT
12801 GAAGTCAGGA TTTCAAGGTC AGCTTGGGCA GCTTAGCGAG AGCCGTCTCT
12851 AGAAAAATAT AAAAGTTAGC TGGGTGTGGT GGCACGTGCC TATAGTCCTA
12901 ACTATTCAGG AAGGTTAGGC GGGAGCACAA CTTGGGTTCC AGGGTTTGAG
12951 GCTCCAGTGA GCTGATCTTG CCACTGCACT ACAGCCTGAG CAACAGAGCA
13001 AGACCCTGTG ACTCCAAAAA CAAACAAACA AACACATTTT GAACCCAAAC
13051 AGATCTGACC CAAGATGCAT GCTCTTATAG ATGCCACCTC CCTGTGTGCT
13101 GGGGCTTCTA CTAAAAACAC AGACAAGCAC AGCCAACCAC AGTCAATCTA
13151 AGGGAAAGAG GAAAGTGTAA CCAAAGCACA AATACATAAA ATATTGCAAA
13201 AATGCTATTT AAAGAAAAAA AAGAGAAGAG AGGCTCTGAG GTTGTACTAA
13251 CAGAGAATGG CCTTGGCTAA TCCAGGAAGA CTTCCTGAAA GAGGTTGTTT
13301 TTTCCCCAGG TCTGCTTTTG ACATCTCTCT TTTCACAGTG CATCTGGGTA
13351 GTGAGCTTCC TCTCCTCCTT CTTCCTCAGC CTGCCCTATG GTGTGGCAGT
13401 GGGTGTCGCC TTCTCCGTCC TGGTCGTGGT CTTCCAGACT CAGTTGTAAG
13451 TGATAGCTTC CGCCCTCCTA GGCCCACGAT CGGTTCCCTG GGCCAGCCCG
13501 CAAAGGGCTT CCATGCCACG GCCTGGCTTA GTCCACTGTA CCTTCCACCT
13551 CTGGGCCTGG CACTGGAGGT GCTGCCAGGC CCAAAGAGAG CCCAACCCAG
```

FIGURE 3D

```
13601 CCAGGACTGT GGGCACAGTC TGGGCTGTTT GACTTCCCAT ATCTTGAAAA
13651 CCCCAGAGAA AGCCAGCATA CTCTTGCTGG GGATGGCTGG GGAGAGGGCA
13701 GTGGCAGAGA AAGGAGGGCA AGGGCAGGTG GTGAGATTCA ACATCCTTCC
13751 AAAGACATTG CCAGAACCCC AAACCAAATG GGACCCCACC CCAGGAGAGC
13801 GCCAGGGTGG AAGACAGAAG CTGTGTTCTA CACACTGGGA GTATTACAGA
13851 GAAGGGGTCT TGGCCAAGGC AGGGAGTACG CTGAATGTTG GGGGAATCCT
13901 ATCTTCTCTT CTTGAGAACT CAGAACAAGG AAATGATGAC TTCAGGGCGA
13951 CTCCCACCAC TTCTCCCACC ACTTCTCTCC CCTGCCCTGT GGTCTGGGAG
14001 CTATGTCAAG GACCTGCCTG TCATCCTCAT AGTTATAGGA GGCCACAGGC
14051 CACCAGACAT GTGTCTCCAG TGCAAAAAGA CAGACACAGC AAGTCTGGGG
14101 GTGAGGACAG GACCCCATCC TACCTTGGCT CTGCCCCCGC CCCAGCAGGG
14151 GCACCCTTCC AGGCCCATGT GCCATTAGCA TTCTCTTATG TTTTTCTCTT
14201 CCTGCTTCAT CCAGTCGAAA TGGCTATGCA CTGGCCCAGG TCATGGACAC
14251 TGACATTTAT GTGAATCCCA AGACCTATAA TAGGGTAGGT AATTCAAGCT
14301 TATGACCTCC TTCTTTTGCT CTGCACCACC CCAAGAAGAG GTTGCTTTTT
14351 AAAGCCAATA AAGACATTTC TGCAACTTGA GCTCAGTCTC CCTGTCACAG
14401 GCCCAGGATA TCCAGGGGAT TAAAATCATC ACGTACTGCT CCCCTCTCTA
14451 CTTTGCCAAC TCAGAGATCT TCAGGCAAAA GGTCATCGCC AAGGTAAGGC
14501 TCAGTCCCTG GCGACCAGAG GCTCTGGACA GAGAGTGGCC GGAAAATGGA
14551 AGCAGAAGGG CGGTGGGAGC TGAGAATAGG CCACTCCCAT AGAGGGTGGA
14601 GGTCAAGATT GCTGTTGGCT CTCTCCCTGC AGACAGGCAT GGACCCCCAG
14651 AAAGTATTAC TAGCCAAGCA AAAATACCTC AAGAAGCAGG AGAAGCGGAG
14701 AATGAGGCCC ACACAACAGA GGAGGTCTCT ATTCATGAAA ACCAAGGTGA
14751 ATGAAGGCCA GAAGCAGCCC CGTGCCCTGC TCTCCTGCCC ATTCTGATAC
14801 TGCCCCCTGT TACTCATGGT ACCCTGGGGG CCCCGCTTCC CACCCTGACA
14851 GGCAAAGACA GAAAGTCTCT GGGAACACTG CCTGGTGGCC GCTGGGCATT
14901 TTTCTTCTTT TTTTTCTTTT TCTTTTTAGA GATGGAATTT TGCTCTTGTC
14951 ACCCAGGCTT GAGTGCAATG GCGTTATCTT GGCTCACTGC AACCTCCACC
15001 TCTGGGGTTC AAGCGATTCT CCTGCCTTAG CCTCCCAAGT CGCTGAGATT
15051 ACAGGTGCCA CCACACCCAG CTAATTTTTG TATTTTTAGT AGATATTGGG
15101 TTTCACCATG TTGGCCAGGC TGGTGTCAAA CTCCTGACCT CAGGTGATCC
15151 ACCTACCTTA GCCTTCCAAA GTGCTGGGAT TACAAGCCTG AGCCACTGCG
15201 CCCAGCCTGG GCATTTTTCT TCTTGGATGA GGTGCTACCA TCTCCCAGGG
15251 AAGCCACTGA ACCCCCAAGG CCCTTCTCCA TTTTCTGGCT AAGATAGGAC
15301 ATGGCCCATG GACTTTTGAA CAACCCAGAG GGGGAACAGC AGTGAATTTC
15351 CTGGGGAACC CAGGCAGCCC AGGGCTAGCA AGGCTGGGGT GGCCATGGCA
15401 GTAATCCTTG TAATCCCAGC ACTTTAGGAG GCCGAGATGG GAGAATCACT
15451 CTCATGAGTT CAGGAGTTCG AGACCAGCCT GCCCAACGTG GCGAAACGCT
15501 GTCTCTACTA AAAATACACA AAAATTAGCC AGGCGTGGTG GTGGGCACCT
15551 GTAATCCCAG CTACTCAGGA GGCTGAGGCA CGAGAATCAC TTGAACCCGG
15601 GAGGCAGAGG TTGCAGTGAG CCGAGATAGT GCCACTGCAC TCCAGCCTAG
15651 GCAACAGAGG GAGACTCTGT CTCAAGAAAT AAAGGAGCTC AGTGTCCCCG
15701 GAGGGGCTTT CTCCCAGAGA GAGTGGGCTT GAGGCTTCAG TGCCTCTCTT
15751 GGCTGGGTCC TCTGACTTTG TCTGGGTTGT AGGAGACCAA GTTTGCAGGC
15801 CCTGCCTAAG AAAGGGCTTT GGGAGAGGCC TCTCTGGTGG AGCTTTCAGG
15851 GTCTGTGTTC ACCATCACCG AGGCGAGTTA TTCCCCTACA CCTACACCCT
15901 CCATGCCCCT GCTTCAGTCA CAGCAAGGTC TGGCTCAGTC TGGTGGTCCC
15951 TGACTCTGCC CACTGTCCCC ACCCTTCCAG ACTGTCTCCC TGCAGGAGCT
16001 GCAGCAGGAC TTTGAGAATG CGCCCCCCAC CGACCCCAAC AACAACCAGA
16051 CCCCGGCTAA CGGCACCAGC GTGTCCTATA TCACCTTCAG CCCTGACAGC
16101 TCCTCACCTG CCCAGAGTGA GCCACCAGCC TCCGCTGAGG CCCCCGGCGA
16151 GCCCAGTGAC ATGCTGGCCA GCGTCCCACC CTTCGTCACC TTCCACACCC
16201 TCATCCTGGA CATGAGTGGA GTCAGCTTCG TGGACTTGAT GGGCATCAAG
16251 GCCCTGGCCA AGGTGAGGCC CTCGGGGACA GCAAGCACCA CCCACTCCAC
16301 CCCCTCCGCT CTGCTCTCCA CATTCCCTTT CCTGGGAGCC CTCATTTCAG
16351 GAAGCTGAGG GAGGAAGCTC ACTGGGGAGA CTAACAGCTC CTAGGAATCC
16401 CTCCTTTCCC CAGACGCCAC CAGGTTGAGA CATTCTCCAC AGAGCAGGCC
16451 CAGACGGCCC ATGACAATGA GTGGCGGGAC AAGTCTACCA GAGTTTCAGG
16501 CCCCTGTGCT CCCAACACCC CCAGCAGTGG CCATCCCAAG TCCCTCTCAG
16551 CCATCAGGAA CCCACCCAGG TTCTCTGAGG AGGGTCCAGT TTGGCTCCTG
16601 GTTCATGATC TGCTGCCCTT GTCCCTCATT CACCAGCCAC CCTAGGACAG
16651 GAGAAGAAAT AATACCAGTG CCCCACACCA TCAGGCCAAA CAGAGAGCCC
16701 ACGGGACACC TTGAATGAAT GTATCCATCT GATAACTTTC CAGCAGCCAC
16751 CGCCAATGGC GGGAGTCAGC AAACCTCAGA GCTGGCTCAG ATAGAGGCAA
16801 GCCAGGGGAA CAATGGGCAC AGAGAGTGTT CGGACTGCCT TCACCATCAA
16851 CCAGGCCAG GGCAGGCCCC ATACCCAGCC TTGGGCCTCA GCCGGCTTCC
16901 TTAGCCAGGA TCTGGAGTCC AGGCCAGCCT TGGCTGAAGC TCTAGACTCC
16951 CTGAGCCTCC ATCCTCCCCT GCAGCTTCTG TCTGAAGCCA CAAAGAAGTC
```

FIGURE 3E

```
17001 TGAGAATCTA AGCTACTGAA AGAAAAGATC AGCCGGGCGT GGTGGCTCAC
17051 TCCTGTAATC CCAGCACTTT GGGAGGCCAA GGCAGGTGGA TCACAAGGTC
17101 AGGAGTTCAA GACCAGCCTG GCCAACATGG TGAAACCCCG CCTCTACTAA
17151 AAATACAAAA ATTAGCCAGG TGTGGTGACG GGCCCCTGTA GTCCCAGCTA
17201 CTCGGTAGGC TGAGGCAGAG AATTGCTTGA ACCCAGGAGG CGGAGGTTGC
17251 AGTGAGCCAA GATCGCGCCA CTGCACTCCA GCCTGGGCAA CAGAGTGAAA
17301 CTCCATCTCA AAAGAAAAAA AAAGAAAATA TCTAGCCCCA CAAGAAGGGG
17351 CCATGGTGAC TTTAAGTGCC CGCCACGTTG GCAAAAGTCC ATTTCCGCTC
17401 CACTTCCCAG AGAAACCGTC AGCCAACACT CCAGGGAGAA GTGGTGTGCT
17451 TTGCTGCTAT TTTTGTCTTT GGCTGCTGGG CTCTCAGGGT TGCTTATTTG
17501 TTTGGCTTCC CCTCTGAAGT ACGTTTTGTG AATCACTTTT GAGACCCACT
17551 CAGAACATTC CTTTCCTTTT GCCTCCCTAC CCCAACAACA CTTCTAGCTG
17601 AGCTCCACCT ATGGGAAGAT CGGCGTGAAG GTCTTCTTGG TGAACATCCA
17651 TGGTAAGAGA AAGAGGACAT TTAGGGACTG AAAGACTGGC AAGGAGTGTG
17701 GGGTAGGAAC AGGTTGGTGG GGTCTGAATA GTGAGGAGGT TGGAAACGAG
17751 AGCACCCAGC TATCCCCCAC AAGCTGCTGC CTGCTCATAA AAGCTTCAGG
17801 TACAAGTCCA AAGAGACTGG TCAGATTGCA TAAACATCCT AGGGGCCTTA
17851 GTGACAGAGT GGGGGTGAGG AGGTCATGGA GTTACAGAAG GACAGCTAGG
17901 ATTCTAATCT ACCCCATAAC TAATTTGCCA CGTATCCTTG GCCGAGTCAC
17951 TTTATCTCTC AAGGGATCTA TTTCTACCTA TGTAAAACGA GAGGGTTGAC
18001 TAGATGGATT TGGGGATCCT CTCCCAATCA GAAACTCTGT GAATCGATAT
18051 AGGCATAGAG CACACGGTAC CCTAATTCCC CAGGGAACAT ATAAATATGC
18101 AGTTTTGTAG GCATACACGG TCCAAAGGGT GCATATACAC AGCCTCAAGG
18151 ACGTGGCCAC AGGGCAGCAG ACATTTACAT GACTAGCATG TACGCAAAGT
18201 GCAGAGATGT GGGAGCAAGT GCACACAGAC ACACAGGAGA ATGTGAAGGG
18251 GCACATACAC ACACACCCAG CTCCCTGCAC TGGGTCAGAC CCCCTCCAGC
18301 AGGGCTGCAG TTCCCAAGCT CCGCATGGCC ACGTTCGGGG AGAGAATCTG
18351 CAGTGGCAAT GACCTGCTAT GATATGTTCT GGAGTTAGAA GCAGTGGATT
18401 CTCCCCAACC TCACTGGACA CCCCCTTAGG AAACCATCTC TAGGATTAAG
18451 AGTAATCCAC ACAAACTTCC AATGCCACAC ATTGGAAGTT GCTGGAAAGG
18501 TCTGGGAAAA CAAGAGGAAG GATGGGTCCT TGGGGGATAG AACTGGCAGC
18551 GGCCTCTTCA AGGATGGCTT AGGCTTTTCC ACTCGAATCA CCACAAAGTA
18601 CTGACTCCCT AAATCAAACT GCTTCCTTCT GCTCTGGGTT GAAACTTCAG
18651 CATCCTCAAG TTCATGTTGC CCTCTGCCGT CCAGAACTGA TATTGCACTG
18701 CCAATGCCAT GGCCCTCAGA TACAGCAAGA GCTGGGACCT CAGGCCTCTC
18751 CCATCCCTGC TCTGGTCTCA CTATCTTCCC CACCCCACAC TCCAATCCAC
18801 AATGGCTGTT ATCTTTCTGA AGGTGATCTT TTCTCCTTCT AGCCCAGGTG
18851 TACAATGACA TTAGCCATGG AGGCGTCTTT GAGGATGGGA GTCTAGAATG
18901 CAAGCACGTC TTTCCCAGCA TACATGACGC AGTCCTCTTT GCCCAGGCAA
18951 ATGCTAGAGA CGTGACCCCA GGACACAACT TCCAAGGGGT AAGGTTCTTG
19001 CACCTGGGGA ATCCTAGGCT CCAAGGCACT GAAATAGCAG GACCAAGAGG
19051 CATTATTAGA AAGAACACAG GAGAAGGTTT AAGTTCCAAT ATCAAGTCTG
19101 CCATTTCAGT TTTCTGAATC TGTTTCCTTA TCTATAGAAT GAGCACCATC
19151 AACTAACATT ACCTACCTCT CTGCATTTTT CTTTTATTTT GTTTTAGGGT
19201 TAAATGATAA TTACATCTTT TGTGTCACTT GAAAGCACTT TGTGTATTGT
19251 AAAAATTCTT TATCAATATA AGTTTTCTGG TTGCACAAAC ACCCAAAGCA
19301 TAGTAGAGCA GGCCCACTCT GCTGGCATCG TTCCCTGCCT CCTCCTCATC
19351 TCTTTCTAAA GGGGGCTTTC GGGAAGGGAG GGGAGGGGAG TAAGCCTACC
19401 CATTTTAACT TACCGGAGCT TAGAGATTTC AGGCTGGTGA GGGATAAAGA
19451 GATTGGGTCT GAGTTTTGTC TCAGCTTTTT GACATTTAAT TTACTAGCTC
19501 AGTAAGTCAT ACAAATGGGA TACAAATAAC ACCATCTAAA ACTCCAGAAG
19551 ACTGGGGAGT CAGAAAAATC CTACCTCCTT GGGGTCCCTG CCCAGATCCC
19601 CAGTCATCTC TAGCCCTCAG GGTCCCCTCC CAGCTCAGCT CCTGCCCTTG
19651 GCCTCCCAAG ACTCTTGTTG TGCCCCAGCC CTGGGTAAAA ACCTCCCCTG
19701 CCCTCTGTGG GTCATAAGAA AGGCTTTTCT GGCCCTAGAG CAATGATTTG
19751 CTCTTTGCCT TAAGAGACTG ATGAAGGTGA AACCATCTGT TCTAAGTGCT
19801 GAAAGACTGC CCAGGAACAC ACAGGGCGCT GGCTCCTGCC CTCCATGCCT
19851 AGAGGGAAAC CCTGGGGAAA CAACGGGCTT TCCTGCTTCG TGAAATTTGT
19901 CCGCAGAGCA AAGAGGGAGA TTCTGGAGGA AGCTGCATTA GTTGTTAGTG
19951 CCCTAATCAT GTTCAGCTAC TCTAGTTGGT ATGTATACTT GATTAGTCAT
20001 AGCACTTATA AATAATTTAT ATTTTATATA ATATATACTT ACATATTATA
20051 GACCATTCAC AGATACAAAT CACACACATA AACACACACC TTTTCAACAG
20101 CATTGTGAGG GACAAAGCAG GCAAAGTGAG GCTGGTTATC AGACTTTAAC
20151 AGATTAGAAA ATATATTCCC AGGAGGACAG GAATTCCCCA AGGTCAGGCA
20201 GCTAGCCAAT AGTTTTTCTA AGCTGAGTAA AACCTTCCCT GCCTCTAACG
20251 GCCCACAAAG GAGGGAAGAC CGCGATACAC ACCTGTCTGG TATAAGGGGG
20301 AAGACCACAG CCGTGCTGTT TTTGTGAGGC AGGTAAGGGA AGGGGCAAGA
20351 GGATAAGTCA TGTGTCAGGA AGCAGCGTCC AACCAGAGCC GGCCACCTGT
```

FIGURE 3F

```
20401 CCCTTTTCCT GCCACCATGC ACCAACTTTG CTGTTCAGTC ACTGAAGCTC
20451 ATTCTGCACT GGCTTCCTCC CTTCCAGGCT CCAGGGGATG CTGAGCTCTC
20501 CTTGTACGAC TCAGAGGAGG ACATTCGCAG CTACTGGGAC TTAGAGCAGG
20551 TGAGCTGAGG GAAGGGGCTG TGAGGGTGGG AGCAGGGCGA AGAGGGGAAG
20601 GATGGGGTCG CTGTCAAATA CAAGGCGTTC ACTCAGCTGT CTCACCTCCA
20651 GCCCAGAGCA GTCACATTCA AGGCCACAAA GATTTGTGGT CATCTTTGTT
20701 TTTTTTCTTT TCCTTTTCTT TTTTTTTTTT TTTTAATTTG AGACAAAGTC
20751 TCACTCTATC ACCCAGACTG GAATGCAGTG GCATGATCTC AGCTCACTGC
20801 AACCTCTGCC TCCCGGGTTC CAGAGGTTCT CCTGCCTCAG CCTCCCGAGT
20851 AGCTGGGACT TCAGGCCTGC GCCCAGCTAA TTTTTGTATT TTTAGTAGAG
20901 ACAGCTTTTC ACCATGTTGG CTGGGCTGGT CTCGAACTTC CGATCTCAAG
20951 CAATCTGCCT GCCTCGGTCT CCTAAGTGCC TGGATTACAG GCATAAGCCA
21001 CGATGCCTGG CCTTTGTTTT CATTCTTCTC ACTCCCTGAA AGGCATCGTG
21051 GGGAGAGGGT GAGTCACTGG ACCAAGTCCT AGAGAACCAG TATCTATTCT
21101 TATTCTCCAA CACATCACCC ACGTGACCCT GAGCAAGCCA CATACACCCT
21151 GGGCCCTAGT TTTTATCATC TGTGAAATTA GGGGAACAT AGGTAATACC
21201 TGTCCCATCC ACCACACAAG ATTGGCAGGG CAGTCACTTG TTCTTTCATT
21251 AATTCAGCAG GTATTTATGG CGTACCTACT GTTTGCCTGA CACAGTTCAG
21301 GATGGGCACA TAGCAGTGAG CAAAACAAAG GCCTCTGCCT TTTAGAAACT
21351 TACGTTATGG TAGAATAGAT GGATTTNNNN NNNNNNNNNN NNNNNNNNNN
21401 NNNNNNNNNN NNNNNNNNNN NNNNNNGTCT ACAAATGAAT TATTATTGCA
21451 TGTGGACAAG CCTTAAGAAC TAAAAAATAT GTGGCTGGGT GCAATGGTTC
21501 ACACCTGTAA TCCCAGCACT TTGGGAGGCT GAGGTGGGCG GACCACCTGA
21551 GGTCAGGAGT TTGAGACCAG CCTGGCCAAC ATGGCGAAAC CCCGTCTCTA
21601 CTAAAAGCAC AAAAATTAGC CAGGCGTAGT GGTGCATGCC TGTAGTCCCA
21651 GCTACTCGGA AGTCTGAGGC ATGAGAATCA CTTGAACCTG GGAGGCAGAT
21701 GTTGCAGTGA GCCGAGATCG TGCCACTGCA CTCCAGCTTG GGTGACAGAG
21751 CTAGACTGTC TCAAAAACAA ACAAACAAAA CAAAACCTAA AAGATATGTG
21801 GATATGAGGG ATCACCATCC CCATAGGGCC CCTGGATTAA CACCACCCCA
21851 CCAATGCCCT GAATTAAAAG AAACCAGATG ACTAGGTTTG GAGAAATCTG
21901 GCTTTGGGTC TATGAGAAGT AGTGTCTCTC TTTGTGCCTC TTCCCATTCT
21951 TTTTGACATT GAGCTCCATG GTGCTCTGAA TCCGTCTCTC ACAGTGCTGA
22001 TGGCAGGTGG GACAGATTAG AAAATAGAGC TGGAGCCACA GAGATTTGGC
22051 AGACTGATTT CGGTGCCCTC TTGGAATCTC CAGCACATTC CAAAAAGCCT
22101 GGATAGGACC AAAATAGCTT ATCAACGTGA GAAAGGACTT CAGAGCTTGT
22151 CTACTGCCAA CCCTCATTTT ACCCAATGAG GAAAGTGAAG CTATTAGGGA
22201 GCGAGGGACA CGTGGAAGGT CACACAGCAC ACAGGAGGTG ATTCACATGT
22251 AGATTTCAGC ACCTGCTCCT GCCACGCTGG ACTGGTTCAC CTCCTAGGCT
22301 GACCCTGCCT CTCCCCTGTT CACACACACT CTCGCACACA CACACACACA
22351 CACACACACA CACAGGTGCT TTGTTCTGGC CAGGGGTTCC TAGGGTCACC
22401 TCTTGGTTGC AGCCACTGTG ACCCCAACTG GTCTAACCTC TCTCTTCCCC
22451 TCCCACTTCC TTCCTGTGGT TCCTGCAGGA GATGTTCGGG AGCATGTTTC
22501 ACGCAGAGAC CCTGACCGCC CTGTGAGGGC TCAGCCAGTC CTCATGCTGC
22551 CTACAGAGTG CCTGGCACTT GGGACTTCCA TAAAGGATGA GCCTGGGGTC
22601 ACAGGGGGTG TCGGGCGGAG GAAAGTGCAT CCCCCAGAGC TTGGGTTCCT
22651 CTCTCCTCTC CCCCTCTCTC CTCCCTTCCT TCCCTCCCCG CATCTCCAGA
22701 GAGAGCCTCT CAGCAGCAGG GGGGTGCTAC CCTTACAGGA GTGAGAGTCT
22751 GGTGAGCCCA CTCTTCACCC GTCAGGCCCT GGCCGCAATG GACAAGCCTC
22801 CTGCTCACTC CACCCCACCC ACCTCTGCCC TGTCCTTGGC AGCTGAAGGA
22851 CACCTTGACT TCCAGCTTTT ACGAGTGAGC CAAAAACAGA AGGACAAGTA
22901 CAACTGTGCT GGCCTGCTGT ACAAGCTTCA AAAAGTGTCC CAGAGCCCAC
22951 ACGGCTCGGT GTCAGATGGT GTCAGGCTGT CACGGACATA GGGATAAACT
23001 TGGTTAGGAC TCTGGCTTGC CTTCCCCAGC TGCCTCAACT CTGTCTCTGG
23051 CAGCTCTGCA CCCAGGGACC ATGTGCTCTC CACACCCAGG AGTCTAGGCC
23101 TTGGTAACTA TGCGCCCCCC GTCCATCATC CCCAAGGCTG CCCAAACCAC
23151 CACTGCTGTC AGCAAGCACA TCAGACTCTA GCCTGGACAG TGGCCAGGAC
23201 CGTCGAGACC ACCAGAGCTA CCTCCCCGGG GACAGCCCAC TAAGGTTCTG
23251 CCTCAGCCTC CTGAAACATC ACTGCCCTCA GAGGCTGCTC CCTTCCCCTG
23301 GAGGCTGGCT AGAAACCCCA AAGAGGGGGA TGGGTAGCTG GCAGAATCAT
23351 CTGGCATCCT AGTAATAGAT ACCAGTTATT CTGCACAAAA CTTTTGGGAA
23401 TTCCTCTTTG CACCCAGAGA CTCAGAGGGG AAGAGGGTGC TAGTACCAAC
23451 ACAGGGAAAA CGGATGGGAC CTGGGCCCAG ACAGTCCCCC TTGACCCCAG
23501 GGCCCATCAG GGAAATGCCT CCCTTTGGTA AATCTGCCTT ATCCTTCTTT
23551 ACCTGGCAAA GAGCCAATCA TGTTAACTCT TCCTTATCAG CCTGTGGCCC
23601 AGAGACACAA TGGGGTCCTT CTGTAGGCAA AGGTGGAAGT CCTCCAGGGA
23651 TCCGCTACAT CCCCTAACTG CATGCAGATG TGGAAAGGGG CTGATCCAGA
23701 TTGGGTCTTC CTGCACAGGA AGACTCTTTA ACACCCTTAG GACCTCAGGC
23751 CATCTTCTCC TATGAAGATG AAAAATAGGGG TTAAGTTTTC CATATGTACA
```

FIGURE 3G

```
23801 AGGAGGTATT GAGAGGAACC CTACTGTTGA CTTGAAAATA AATAGGTTCC
23851 ATGTGTAAGT GTTTTGTAAA ATTTCAGTGG AAATGCACAG AAAATCTTCT
23901 GGCCTCTCAT CACTGCTTTT CTCAAGCTTC TTCAGCTTAA CAACCCCTTC
23951 CCTAACAGGT TGGGCTGGCC CAGCCTAGGA AAACATCCCC ATTTCTAACT
24001 TCAGCCAGAC CTGCGTTGTG TGTCTGTGTG TTGAGTGAGC TGGTCAGCTA
24051 ACAAGTCTTC TTAGAGTTAA AGGAGGGGGT GCTGGCCAAG AGCCAACACA
24101 TTCTTGGCCC AGGAGCATTG CTTTTCTGTG AATTCATTAT GCCATCTGGC
24151 TGCCAATGGA ACTCAAAACT TGGAAGGCGA AGGACAATGT TATCTGGGAT
24201 TCACCGTGCA CAGCACCCGA AGTGCCAAAT TCCAGGAGGA CAAGAGCCTT
24251 AGCCAATGAC AACTCACTCT CCCCTACTCC ACCTCCTTCC AAGTCCAGCT
24301 CAGGCCCAGG AGGTGGGAGA AGGTCACAGA GCCTCAGGAA TTTCCAAGTC
24351 AGAGTCCCCT TTGAACCAAG TATCTAGATC CCCTGAGGAC TTGATGAAGT
24401 GATCCTTAAC CCCCAAGTAA TCATTAACCC CCAGACCAGC CTCAGAACTG
24451 AAGGAGATTG TTGACCCAGT GACCTGGAGT TGAGGCTCAG GGAGAGATCT
24501 GCCACATGTC TGAGGGTTGC AGAGCC (SEQ ID NO:3)
```

FEATURES:
Start: 1997
Exon: 1997-2121
Intron: 2122-4732
Exon: 4733-4872
Intron: 4873-5004
Exon: 5005-5115
Intron: 5116-5781
Exon: 5782-5957
Intron: 5958-7770
Exon: 7771-7935
Intron: 7936-8470
Exon: 8471-8623
Intron: 8624-8917
Exon: 8918-9000
Intron: 9001-9777
Exon: 9778-9925
Intron: 9926-10221
Exon: 10222-10335
Intron: 10336-10539
Exon: 10540-10617
Intron: 10618-11197
Exon: 11198-11293
Intron: 11294-13338
Exon: 13339-13445
Intron: 13446-14214
Exon: 14215-14284
Intron: 14285-14400
Exon: 14401-14493
Intron: 14494-15980
Exon: 15981-16262
Intron: 16263-17597
Exon: 17598-17652
Intron: 17653-18842
Exon: 18843-18988
Intron: 18989-20477
Exon: 20478-20549
Intron: 20550-22478
Exon: 22479-22523
Stop: 22524

CHROMOSOME MAP POSITION:
Chromosome 1

ALLELIC VARIANTS (SNPs):

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
48	C	G	Beyond ORF(5')			
132	G	A	Beyond ORF(5')			
724	A	C	Beyond ORF(5')			
1558	C	G	Beyond ORF(5')			

FIGURE 3H

1577	A	G		Beyond ORF(5')			
2487	C	A		Intron			
2634	T	C		Intron			
4352	A	G		Intron			
5157	A	C		Intron			
5658	A	T		Intron			
5945	T	C		Exon	180	T	T
6281	C	T		Intron			
6452	G	C		Intron			
6610	T	G		Intron			
7247	T	C		Intron			
7360	A	G		Intron			
7644	A	T		Intron			
8127	A	C		Intron			
8317	G	A		Intron			
9079	G	A		Intron			
9537	G	T		Intron			
12302	C	G		Intron			
12354	C	T		Intron			
12487	C	T		Intron			
13198	-	A		Intron			
13257	A	G		Intron			
14541	G	A		Intron			
14545	A	G		Intron			
15041	C	A		Intron			
15053	A	C		Intron			
15065	A	G		Intron			
15108	A	C		Intron			
16274	-	G		Intron			
17424	C	T		Intron			
17627	G	A		Exon	657	V	V
18427	T	C		Intron			
18813	C	G		Intron			
19035	T	C		Intron			
19182	T	C		Intron			
19508	-	G	C	Intron			
19571	T	G	C	Intron			
20147	T	G		Intron			
20180	G	A		Intron			
20584	A	T		Intron			
20717	T	C		Intron			
20894	A	G		Intron			
21787	-	A	C	Intron			
22264	T	C		Intron			
22338	-	C	A	Intron			
23363	T	C		Beyond ORF(3')			
23688	G	A		Beyond ORF(3')			
24210	A	C		Beyond ORF(3')			

Context:

DNA
Position

48
CTGGGTTCCTATGTGGGGAGGTCATGCTCCCCACTCATTGAGCCCCC
[C,G]
CAGGCAAACCACCTGGACAGCCAGACCCATGCAGACTCTGGAGCAGGTGGAGAGGAAGAG
TGAGACCACCCCGCCTCACGGGCGGTGAAGGGCCGGCAGCCTCTGAATAGTCTCTGCTAG
GAGGTAGAAAGCACCCTCCCATCTTAATCATAGTAATCATCGCCACTACCATTTACTGGG
TGCCTATAAAAGGCCAGCCTCTTCATACACATGATCTCACTGAATCCTCATAGCATCTGC
CTGCGACTGTTATTATCCCCATTTACAGATGAAGAAACTGAATCTTTGAACCCAGGTCAT

132
CTGGGTTCCTATGTGGGGAGGTCATGCTCCCCACTCATTGAGCCCCCCCAGGCAAACCAC
CTGGACAGCCAGACCCATGCAGACTCTGGAGCAGGTGGAGAGGAAGAGTGAGACCACCCC
GCCTCACGGGC
[G,A]
GTGAAGGGCCGGCAGCCTCTGAATAGTCTCTGCTAGGAGGTAGAAAGCACCCTCCCATCT
TAATCATAGTAATCATCGCCACTACCATTTACTGGGTGCCTATAAAAGGCCAGCCTCTTC
ATACACATGATCTCACTGAATCCTCATAGCATCTGCCTGCGACTGTTATTATCCCCATTT

FIGURE 3I

ACAGATGAAGAAACTGAATCTTTGAACCCAGGTCATCTGGCTCTCAAACTTGTGCTGTTT
TCCCTAAGCCACCCGGTCTCTCATTTCTCCCACTGAAATGTCTCACATGCCATTGCCCTT

724   ATTGCCCTTACTCATTTCTGCCCATGTCTCCTCCAAAACACCATTTATCAATTCGCTCAA
CAAGTATGTGTTGAGTACACACTAAGGGCCAGGCGAGGGGCTGGGCACAGGCGCTGGGGG
TAGGTTCATTCTCCCACCTTCGCTTCTGCTGGGTATCACCTGTGGGGTCTTGCCGGGCAT
CCCACCCTCACCTGTAGTTCAAGTGGACCTTGGGATCCCAAGACCAAATGAATGGAATGC
ACCAGCCCAGCCTTCACCAACTTGAGCACAATCTTATTCATAATAGAAACTCACATTTGC
[A,C]
TCACACTTTACATTTTACACAACCCCTTCTTATCCATTAACTCATTTGATCTTCACAACA
ACCCTGTGAGATATGTCTGTTACTCCCACTTTAGTGATACAGAATCTGAGGTTTGAAAAG
TAATGCTGACCATTCTGCCTCATTAATAAAAGCAGGATTAACCCAGGCTCCTGGACCCTT
CCACAAAAGGCATTAAGCAACCTGCTCCCCTCTGACAACCTCCCCTGTCACCCAGGCTCT
CCTCTGGGAAGTTGGGGGCATCTCTAGCCCCCAAGTAGTTACTCATTTTCAACCCCATCT

1558   TCAGCTCTGCCCATCTCAGCTCCTGGAACGTCAGCCAGGTTGCGCAAAAAGTGAGGAGGA
GAGGAGCGGCAGTACACAAGGGTGGGGGAAAGATTAGGCACAGGAAGCCGTGGGAGAGAG
AGCCGGCAGGTGGACCATCCTGGTTTCCCCACACACACCATTGTCCCCCTGGGAAACCTG
TTGGTGAAGTTCTAGATGTCTTATCCAAGAAGGGTCCTCTTGAGGTCATCTCAGCTATCC
CCCTGCCTCTAGGCAAGCTGTTTTCTGTTTCTTCCAAGCTGACTGGCTGAATGGTAGGAG
[C,G]
CTTTCTGCCAGGGAAACTAAGGTCTGGGAAGGGAGTATGGCTTGTGGGGACACCAGGGGT
CAGGGGAGGGGAGGGTCCACCTGCTGAATCAAGTGGGGCCTCCTGCCCTCGTGATTCCCC
TTTGCCTGGTGCTCAGTGGGGGTGATGGTGACGCCACAGGTGTGGAGTGCCAGCCACGTG
CTGAGCGCCAAGCAAAACAGCCAGGGTGAGTCTATGCATCATCAGTGCCTGGGAAGGAAG
GCCACTGCGAGCAGGGAGTCTGACGGAAAAAACTTGACAGAGGGAAGGGAGGCACCTTGCT

1577   CTCCTGGAACGTCAGCCAGGTTGCGCAAAAAGTGAGGAGGAGAGGAGCGGCAGTACACAA
GGGTGGGGGAAAGATTAGGCACAGGAAGCCGTGGGAGAGAGAGCCGGCAGGTGGACCATC
CTGGTTTCCCCACACACACCATTGTCCCCCTGGGAAACCTGTTGGTGAAGTTCTAGATGT
CTTATCCAAGAAGGGTCCTCTTGAGGTCATCTCAGCTATCCCCCTGCCTCTAGGCAAGCT
GTTTTCTGTTTCTTCCAAGCTGACTGGCTGAATGGTAGGAGCCTTTCTGCCAGGGAAACT
[A,G]
AGGTCTGGGAAGGGAGTATGGCTTGTGGGGACACCAGGGGTCAGGGGAGGGGAGGGTCCA
CCTGCTGAATCAAGTGGGGCCTCCTGCCCTCGTGATTCCCCTTTGCCTGGTGCTCAGTGG
GGGTGATGGTGACGCCACAGGTGTGGAGTGCCAGCCACGTGCTGAGCGCCAAGCAAAACA
GCCAGGGTGAGTCTATGCATCATCAGTGCCTGGGAAGGAAGGCCACTGCGAGCAGGGAGT
CTGACGGAAAAACTTGACAGAGGGAAGGGAGGCACCTTGCTTTATCGGGGCGGGGAAGGC

2487   ACACGGCTTCTGCACTGGTATCCCTAAGATGGGGTTAAGGGAAGCCCTGGGGAAGTGAGG
TTCTGAATGATGAATTTAAGATCCTACAACCTCATCTGTACTGAGACCCCCAGGGAGGAT
GGGGAGCAGGAGCAAGAACCATCCAGAAGGGTTATATGGCATTCCCAAACCCCTGCATGG
CATCTCCCATATTCTCAATTCACCCGGGTCTCTCTGGGTTTGTTAAGGCATGGTAGATGA
GCATCTACGTTATGGAGGGGTGGGGAGCATCAGAGCCCTTACTCCATGCCCTGTTCCCTC
[C,A]
TTACAAAAAATACCTGAAGTTACCATCACCCCAGGTTCTTTGTCCTTTCCCTCCCGGATG
TTCCTTCCTCCACTTGGTCCAGAGAATGCCAAAAGGAGGCCCTAAATTTCTGAACTTTCC
TGAGGGGACCTACCAGGGTGTAGTCCTACCAGCGCCCAGGGTCTTTCCACTCTCATCTCC
CTGGAAATGCGATGGTGGGTATGAAACCTTGTCCCTAAGTAGGCGCTACACAAGGTGATC
CATACCCACACCCCAGGAGGCTGGGGCTGCGGGTGTCACCCTCCCCATTCCCAGACTCCT

2634   AGGGTTATATGGCATTCCCAAACCCCTGCATGGCATCTCCCATATTCTCAATTCACCCGG
GTCTCTCTGGGTTTGTTAAGGCATGGTAGATGAGCATCTACGTTATGGAGGGGTGGGGAG
CATCAGAGCCCTTACTCCATGCCCTGTTCCCTCCTTACAAAAAATACCTGAAGTTACCAT
CACCCCAGGTTCTTTGTCCTTTCCCTCCCGGATGTTCCTTCCTCCACTTGGTCCAGAGAA
TGCCAAAAGGAGGCCCTAAATTTCTGAACTTTCCTGAGGGGACCTACCAGGGTGTAGTCC
[T,C]
ACCAGCGCCCAGGGTCTTTCCACTCTCATCTCCCTGGAAATGCGATGGTGGGTATGAAAC
CTTGTCCCTAAGTAGGCGCTACACAAGGTGATCCATACCCACACCCCAGGAGGCTGGGGC
TGCGGGTGTCACCCTCCCCATTCCCAGACTCCTGGCAGACCTCCTCTGGCCCAGCTATAG
GCCAACTCACTCTCCCTCACTCCCTTGGGGAAACGGCTGATTCAGTTACCTGGATTGAGG
TCACTGGCAATGGCTGAAGTGGAGACGCAGGTGGAACTGGTTCAGGCCGGGGGAATCACC

4352   ATTGGAGTTACCACACATAAAGGATAGTGAGTCAGCAGAGTGCACCCTGCAGGAACAATA
GAGCCTTCCTTTTCAAGGAAGTTCTAAGAAAAATGGCAGCAGGCAGGCCCCACTCGGGTG
TATTCACTCATTCATTTATTCAACAAATATTTACTAAGTGCCCCTGTGCAAGGCTCGAGG
TGTACAAAGATGAACAGGAGAGCTAGACTTCTTGCCATGCGTGGTGGGGTTTGCTGCCTA
GTGGGAGAGACAGACAAAAAGCAAGGAATGCACACACAGGATGCACACACAGCGGCAGGA

FIGURE 3J

[A,G]
CCAAGGTGCAGTTACCCAGGCCTGGGATCAGACAGACAGGACTCAGAGGAGACTTTCCCA
GAGAAAAGCCATCTGAGCCAAGGGATGGATCTGATACCTCCGAAGGCTGAGCCACCATAA
CACTCATACCTTTAAGCCAAGTCTTATAAACTCCCCAGGTAAGCAGCTGGCAGTCAGAAG
ACCTCCAGCTAATGCCCAGGACAAGTTGATGAGCTCTCAAGAAAAAGTTCCTGCCTTTTC
TTCTCAATATCCCTGGCACACAGTTCAGTGAATTTTGAATGAACCAATGAATGAAATGAG

5157   ATCCAGGTCCCACAAGGTGAAGGGGCTCCTTCAGCCAGGCCTGGATTGCCACTCCCCTCA
CCATTCCTCTCCTCATCCCCACTCCATCCCTCTGTGATCCCCATAAGCTAGTCATGCTGC
TGAGCTTCAGTCTCGTTGTCCTCTGCAGGCATGGCATTTGCTCTGCTGGCCAACCTTCCT
GCAGTCAATGGCCTCTACTCCTCCTTCTTCCCCCTCCTGACCTACTTCTTCCTGGGGGGT
GTTCACCAGATGGTGCCAGGTAAGGCCTCTCCCCTCTGGGCAGGCAGGATGACCCAGACC
[A,C]
CAAGGATGGGAGGTGTGGCAAAGGGGCCTCGGGAGATTTTCCATCTGCATTCTCCTGGAG
TTGTTCCTGGTCAGTCCTAGGGGAATGGTCACTGTGAATGTCATTTCCAGGTCCTCGGTG
ACCTTGGAGAAACCACTGAGCCTCTTTGAGTTCAGTTAGCATTACCTGTTCCATCTTCCT
CCTAGGAATGAGAGGAAGACTTAGCAGAACAAGATATACCATATGCTATAACATGCTTAA
ACAGATGTGAGAAATCACCATCTAACTCCCTGGTTGGTCCCAGCCGGCCACTACAGGGAC

5658   TTAGCAGAACAAGATATACCATATGCTATAACATGCTTAAACAGATGTGAGAAATCACCA
TCTAACTCCCTGGTTGGTCCCAGCCGGCCACTACAGGGACATTTGGACTTCTCTGGTGCT
AAGTGAGATGGAGGAAAGCCTGGTCACAAGGGCTGGTTTCTGGTTCAGGCTCTGCTTATA
TTTCTTATTTCTGAGTTCATTTTCTCACGTGTCCTGTATGACAATATTGACCATTGGGGT
AAAAGCACCTTGAAAAGCATAGATCATGGTTAGAGTGAGTGGTTGTTATTATTGTGTTGG
[A,T]
GAAGAGCCTTGGAGGTGCAGGGATCCATCCCCCTGGGGTCGGGAAGCATTCCTGGGCCCC
TTTCTGGTTTCCATCGGTGTGGTTCAAACCTCTGATTTTTGCTGGCTGGGTGGGGCACCA
CAGGTACCTTTGCCGTTATCAGCATCCTGGTGGGTAACATCTGTCTGCAGCTGGCCCCAG
AGTCGAAATTCCAGGTCTTCAACAATGCCACCAATGAGAGCTATGTGGACACAGCAGCCA
TGGAGGCTGAGAGGCTGCACGTGTCAGCTACGCTAGCCTGCCTCACTGCCATCATCCAGG

5945   ATTATTGTGTTGGAGAAGAGCCTTGGAGGTGCAGGGATCCATCCCCCTGGGGTCGGGAAG
CATTCCTGGGCCCCTTTCTGGTTTCCATCGGTGTGGTTCAAACCTCTGATTTTTGCTGGC
TGGGTGGGGCACCACAGGTACCTTTGCCGTTATCAGCATCCTGGTGGGTAACATCTGTCT
GCAGCTGGCCCCAGAGTCGAAATTCCAGGTCTTCAACAATGCCACCAATGAGAGCTATGT
GGACACAGCAGCCATGGAGGCTGAGAGGCTGCACGTGTCAGCTACGCTAGCCTGCCTCAC
[T,C]
GCCATCATCCAGGTGAGGGGGCAGCCCCCAACCCTGCTAGAAGGGCATCAGACCACCCTG
CCCCTCCCTCAAAGCCTTAGCTTTGATGCTAAATCTGATTTAGGGGGGCTGGGTGTGGAGG
CTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGAGGGTGGATCACTTGAGGTCAGG
AGTTTGAGACCACCTTGACCAACGTGATGAAACCCCATCTCTACCAAAAATACAAAAATA
ATCCAGGCTTGGTAGTATGCGCCTGTAGTCCCACCTACTCAGGAGGCTGAGGCAGGAGAA

6281   GCTAGAAGGGCATCAGACCACCCTGCCCCTCCCTCAAAGCCTTAGCTTTGATGCTAAATC
TGATTTAGGGGGCTGGGTGTGGAGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGA
GGAGGGTGGATCACTTGAGGTCAGGAGTTTGAGACCACCTTGACCAACGTGATGAAACCC
CATCTCTACCAAAAATACAAAAATAATCCAGGCTTGGTAGTATGCGCCTGTAGTCCCACC
TACTCAGGAGGCTGAGGCAGGAGAATCACTTGAATCCGGGAGGCAGAGGTTGCAGTGAGC
[C,T]
GAGATCGCGCCACTGCACTCCAGCCTGGGTGACAGAGCGAGACTCCGTCTCAAAAAAAAA
AAAAAAAAAAAAAAAAAAAACCCAAGTTAGGGCTCACCTCCTCCCTCCTCCCCATCCCAGG
GCTAAAGTGAACCTTGAAAATTAACAGTATCTCCTCATCTGCATGTAGCAGGACCATACA
AAAAAACAACAGCTGTACCTGGTTAAACTGTCCTGAGCTTTAAACCTGTAAAAGACTCAC
AGCCTCTCTCCATTATCCCGTGGAGAAACCCAACTCTCTGCCAGCATAGTCTTGCAGACT

6452   ATGAAACCCCATCTCTACCAAAAATACAAAAATAATCCAGGCTTGGTAGTATGCGCCTGT
AGTCCCACCTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAATCCGGGAGGCAGAGGTT
GCAGTGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGTGACAGAGCGAGACTCCGTCT
CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCCAAGTTAGGGCTCACCTCCTCCCTCCTCC
CCATCCCAGGGCTAAAGTGAACCTTGAAAATTAACAGTATCTCCTCATCTGCATGTAGCA
[G,C]
GACCATACAAAAAAACAACAGCTGTACCTGGTTAAACTGTCCTGAGCTTTAAACCTGTAA
AAGACTCACAGCCTCTCTCCATTATCCCGTGGAGAAACCCAACTCTCTGCCAGCATAGTC
TTGCAGACTGCTAATTTTCTCTAACATCCCTCACTCCGCTCCAGCCTCCTCTGCTCCAAG
CCACAGCAGCAGTTGCACAACATAAATTGAGCTTCTGCAAATGGTTGCAAAGGATTCTGC
TAGGTTTTATGAAGGGAAGCACAACATGACAGAATGCAAGAGCAAAACACAGTCCCAGAG

6610   GTGACAGAGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCCAAGTTA

FIGURE 3K

```
        GGGCTCACCTCCTCCCTCCTCCCCATCCCAGGGCTAAAGTGAACCTTGAAAATTAACAGT
        ATCTCCTCATCTGCATGTAGCAGGACCATACAAAAAAACAACAGCTGTACCTGGTTAAAC
        TGTCCTGAGCTTTAAACCTGTAAAAGACTCACAGCCTCTCTCCATTATCCCGTGGAGAAA
        CCCAACTCTCTGCCAGCATAGTCTTGCAGACTGCTAATTTTCTCTAACATCCCTCACTCC
        [T,G]
        CTCCAGCCTCCTCTGCTCCAAGCCACAGCAGCAGTTGCACAACATAAATTGAGCTTCTGC
        AAATGGTTGCAAAGGATTCTGCTAGGTTTTATGAAGGGAAGCACAACATGACAGAATGCA
        AGAGCAAAACACAGTCCCAGAGAGCGCCTTTTCATTCACTCATTCATTCGGTTTTGTGCC
        AAGAACTAGGCTAAACCCTGGGATACAAAGATAAGTAAGAAAGAGGTCCAATTCACAAGT
        TGCTCACAGCCCAGCAGAGGAAGGAGCCATGTCAACAGATAAATTTGTATGCAGTGAGAT

7247    GACACAGAGCAGAGTCACGGAGGACCTCAAAGAGGAGGTGACACTCCACCTCTCTTAAAG
        GATGAGAACTTAACCAGGAACAAGGTATACAGAGGATGGTCCAGGCAGAAGGGAACAGTG
        CCTAAAAACACTGAGGCCTGAGAGAGTGTGATCTGCGCAGGCAAAGTAAGGGGCTTGGTG
        TGGCTGGAGGGTAGAGGGCCCAGAAGAGGATGGAAAAGTAGGCAGGAGCCAGACAATGAG
        ATCTGGGGTCTGTTCTCTGACAGCGACTTTGGGTCTGATTGGCAGTTTATAAGGATCGTT
        [T,C]
        GGGCTACACAATGATGAGTGGGAGGTGGATTAGAATCAAGGCAGGGGACCTGTTGGGAGA
        CTCTGCAGAGGCCCAGGCAGGAATAATGCAGGCGAAGACCAGGTAGAGAAAGAGATGGGG
        CTGGACTTGAAAAGAATGTTTTACCAGGAGCTTGGTGATAGACTGGATGTGGGAGGTAAG
        GGAGGATGACTCTCAAGTTTTTGGTTGGGCAACCAGGTTAATGATGGTGTCATTTACTGA
        GAGAGAAAACACTGGGGGAGGACTAGACTTATTTTACAGATAAGCCAAAGCCAGAGAGGT

7360    AACAGTGCCTAAAAACACTGAGGCCTGAGAGAGTGTGATCTGCGCAGGCAAAGTAAGGGG
        CTTGGTGTGGCTGGAGGGTAGAGGGCCCAGAAGAGGATGGAAAAGTAGGCAGGAGCCAGA
        CAATGAGATCTGGGGTCTGTTCTCTGACAGCGACTTTGGGTCTGATTGGCAGTTTATAAG
        GATCGTTTGGGCTACACAATGATGAGTGGGAGGTGGATTAGAATCAAGGCAGGGGACCTG
        TTGGGAGACTCTGCAGAGGCCCAGGCAGGAATAATGCAGGCGAAGACCAGGTAGAGAAAG
        [A,G]
        GATGGGGCTGGACTTGAAAAGAATGTTTTACCAGGAGCTTGGTGATAGACTGGATGTGGG
        AGGTAAGGGAGGATGACTCTCAAGTTTTTGGTTGGGCAACCAGGTTAATGATGGTGTCAT
        TTACTGAGAGAGAAAACACTGGGGGAGGACTAGACTTATTTTACAGATAAGCCAAAGCCA
        GAGAGGTGATGTGACAGAAAGGCCCATGCTCTAAAGGAGCTGAAGGTCTGATGGCAGCCA
        TGTAGAGCACAGTGAAGGGCAGGTGAAGGTCACAGATGGTCCAATTCCCTCAAGCTACTG

7644    GACCAGGTAGAGAAAGAGATGGGGCTGGACTTGAAAAGAATGTTTTACCAGGAGCTTGGT
        GATAGACTGGATGTGGGAGGTAAGGGAGGATGACTCTCAAGTTTTTGGTTGGGCAACCAG
        GTTAATGATGGTGTCATTTACTGAGAGAGAAAACACTGGGGGAGGACTAGACTTATTTTA
        CAGATAAGCCAAAGCCAGAGAGGTGATGTGACAGAAAGGCCCATGCTCTAAAGGAGCTGA
        AGGTCTGATGGCAGCCATGTAGAGCACAGTGAAGGGCAGGTGAAGGTCACAGATGGTCCA
        [A,T]
        TTCCCTCAAGCTACTGCTACGCTAGGACTGCACGGAGCTCCAGACCTGCGTGTGTGTGGG
        GCGGGTCGTTGGAACTGCTGAACCACATTGGTCTTCCGCCACCAACCACCCTTTTCCTCC
        TCTCAGATGGGTCTGGGCTTCATGCAGTTTGGCTTTGTGGCCATCTACCTCTCCGAGTCC
        TTCATCCGGGGCTTCATGACGGCCGCCGGCCTGCAGATCCTGATTTCGGTGCTCAAGTAC
        ATCTTCGGACTGACCATCCCCTCCTACACAGGCCCAGGGTCCATCGTCTTTGTGAGTCTG

8127    CATCCGGGGCTTCATGACGGCCGCCGGCCTGCAGATCCTGATTTCGGTGCTCAAGTACAT
        CTTCGGACTGACCATCCCCTCCTACACAGGCCCAGGGTCCATCGTCTTTGTGAGTCTGGG
        GATGCACCCCTGCCATTGGAGCAAGGCTCCAGCAGACACATGAGGAGGATGTACTGTTTT
        AAGATGTCGTGAGCTCCTCATTGCAAGGGCTGGCTTAGCTGTTGTTCAGAGAGGATTCTG
        AGGGGGTTTCTGTCTTGGGAGGGTCAAAGTCATGACTCACAGAGGTTCTTGGTAGTTAAT
        [A,C]
        CCTGCAGAAAAGAGCTGTACATTCTCCGCCAGTTCCCCATTCTAGTGCCTCAACCCCTCC
        CTGCCTGGAAAGTCCTGCCTTATGTCTAATCTCCATCCCTCCTCCTTCAGCCCAAACTCT
        TCTAAAGAAAAAGAAAGCATTCCTTTTCTAGCACAAGTTCCCCATGTGCCTTTTGGGAAA
        GGGCGGTGGGCGACGGGACAGGGTTCCTGATCAGGGTTTTAATTCTGTCTTGGTGTGCCT
        CCATTAGCTTTGATGGCATCCCTTCCCTGGGTCAGACACCCAAAGGTGGGGTATTATGGG

8317    GAGCTCCTCATTGCAAGGGCTGGCTTAGCTGTTGTTCAGAGAGGATTCTGAGGGGGTTTC
        TGTCTTGGGAGGGTCAAAGTCATGACTCACAGAGGTTCTTGGTAGTTAATACCTGCAGAA
        AAGAGCTGTACATTCTCCGCCAGTTCCCCATTCTAGTGCCTCAACCCCTCCCTGCCTGGA
        AAGTCCTGCCTTATGTCTAATCTCCATCCCTCCTCCTTCAGCCCAAACTCTTCTAAAGAA
        AAAGAAAGCATTCCTTTTCTAGCACAAGTTCCCCATGTGCCTTTTGGGAAAGGGCGGTGG
        [G,A]
        CGACGGGACAGGGTTCCTGATCAGGGTTTTAATTCTGTCTTGGTGTGCCTCCATTAGCTT
        TGATGGCATCCCTTCCCTGGGTCAGACACCCAAAGGTGGGGTATTATGGGAAGAAGGGGT
        GGGAGCCTGTGAGCATGATGCTCTTTCCCCCAGACCTTCATTGACATTTGCAAAAACCTC
```

FIGURE 3L

```
        CCCCACACCAACATCGCCTCGCTCATCTTCGCTCTCATCAGCGGTGCCTTCCTGGTGCTG
        GTGAAGGAGCTCAATGCTCGCTACATGCACAAGATTCGCTTCCCCATCCCTACAGAGATG

9079    TTCTACTGCTCTAATAATTCCCCCTAAGGAGGCAGGGGAGTGGGATTCAGGGTCCCCAGA
        GAAAAGGGAGACTTGAGAGAGACGCCTGCCCTGGCCCCACCTTAGGGCCAATCCCCATTC
        TCCACTCTGGGGTTTGCAGGTGGTGGTGGCAACAGCTATCTCCGGGGGCTGTAAGATGCC
        CAAAAAGTATCACATGCAGATCGTGGGAGAAATCCAACGCGGGTGAGTCCAGGTGGCCCA
        GAAGCCTGGCCCACCCGCACCTCATGCCCCACTAAGGCCTGAGCTCGGAGAGGGAGACAA
        [G,A]
        ATGAACTCTATGAAAGTGCAGTCGAAACTGTATGACACTGACCATGTATGAATTATTACT
        ATTACCGTTTCCTGAGAAGGGCCGCACAACCAGCCAATGTAGGCTATTTTATGAGAAATG
        AGTCTTAACTGCCACACTCCCCTTATAAATCTCATTCAACTGATGCTGTTAAACAAAGCC
        TCTCTGAACAGCCGCTTGCTGGCTCTTTGCCTTGCTCTAATGCATTGGTTCTTTGTCCAT
        GTAGAAAGGGAACTATTAGGTTCAACCAGATTCATGAAGCATCCACTCTGTGCCAGGCAC

9537    AACTGATGCTGTTAAACAAAGCCTCTCTGAACAGCCGCTTGCTGGCTCTTTGCCTTGCTC
        TAATGCATTGGTTCTTTGTCCATGTAGAAAGGGAACTATTAGGTTCAACCAGATTCATGA
        AGCATCCACTCTGTGCCAGGCACCATGCTGGGCCCTGGGAGGAGAGGGGTGACGCTTGTC
        CTGCAGGGTTGGAACAGGCAAGGGAGGGAAGACCACATAGCACCAAAGGTCTAGGGGTCT
        GTGGACTCGTGAGCATACAGGGTTCAGAATCTGGGAGTTAACAAACGAGGCCCTACCACA
        [G,T]
        ACTGGCCCGGGGGACCTTGGGCAAGTTAGGTTCTCTCAGCCTCAGTTTCCTCCTTTGTAAA
        ACAGGAGTGATGGTCCCTACCCTATGGGGTGGTGCTGAGGATTCAGACTGGATGGGATAA
        CTTAGGCAAAGATCCCGGCACACCATGGGGGCCTGGCTGGTCCCTGTGGGCTGGTGAAGG
        ACTTGGCTGCCCTCCCCACTCACACCCTTGGGTTCTGCCTCCTTCCTGGCTCCTCGGCAG
        GTTCCCCACCCCGGTGTCGCCTGTGGTCTCACAGTGGAAGGACATGATAGGCACAGCCTT

12302   AGCCCCACCATAACCTATGGGAGAGGATTTACTAAACTTTCTTAACGGTGATGAAACCAA
        GGCTCAGAATGGTTAAGTAAATTGTCAAAGGCCACAGAGGTAGGGAGTGGTAGAGTCTGG
        ATTAAAACTCCAAGTCCTGGACTCCAGACCTCTAGGCTGTACTGTCTCATAGGGAAGGCA
        GTCTCACCCACCTAGGGCAGAGAAGAAAATCCTTAAAGCCAGAGAAGTGAGTGGCTCATC
        TGTGGTCACCCAGAGAGACAGTGATGAGGACAGGGAGAAAAATTATACCTCAGTTCCCAG
        [C,G]
        CCAAGGATCTGCTTTGACCATAACCCAACAAGCCCCCGCTATGGTGGTATTTCCTTAGGT
        TCATATGGCGGCTTTTGTTTCCATTTGATCTTCACAGCAATTCTCTACAGGAATCTGGGC
        AGATTTATTTCCTTTAGAGGAATTTCCAGGTCTTAAAATCTATAGGGGGCAACTATCAAA
        ACTTCACCCAATGTTGCCCCCTACCCACACACAAAACCAGGCCCCCAGCCGATCAGAAAG
        CACTGCTGAGCTCCTGTCAGGGCCCACGCAGCTCGCTGTGAGACAGAGAGAGGGAACTCA

12354   GAAACCAAGGCTCAGAATGGTTAAGTAAATTGTCAAAGGCCACAGAGGTAGGGAGTGGTA
        GAGTCTGGATTAAAACTCCAAGTCCTGGACTCCAGACCTCTAGGCTGTACTGTCTCATAG
        GGAAGGCAGTCTCACCCACCTAGGGCAGAGAAGAAAATCCTTAAAGCCAGAGAAGTGAGT
        GGCTCATCTGTGGTCACCCAGAGAGACAGTGATGAGGACAGGGAGAAAAATTATACCTCA
        GTTCCCAGCCCAAGGATCTGCTTTGACCATAACCCAACAAGCCCCCGCTATGGTGGTATT
        [C,T]
        CCTTAGGTTCATATGGCGGCTTTTGTTTCCATTTGATCTTCACAGCAATTCTCTACAGGA
        ATCTGGGCAGATTTATTTCCTTTAGAGGAATTTCCAGGTCTTAAAATCTATAGGGGGCAA
        CTATCAAAACTTCACCCAATGTTGCCCCCTACCCACACACAAAACCAGGCCCCCAGCCGA
        TCAGAAAGCACTGCTGAGCTCCTGTCAGGGCCCACGCAGCTCGCTGTGAGACAGAGAGAG
        GGAACTCACATTTATTGATCACCTACTGAGCATCCATCACTAGGCTAGGACCGTCACATT

12487   ACCCACCTAGGGCAGAGAAGAAAATCCTTAAAGCCAGAGAAGTGAGTGGCTCATCTGTGG
        TCACCCAGAGAGACAGTGATGAGGACAGGGAGAAAAATTATACCTCAGTTCCCAGCCCAA
        GGATCTGCTTTGACCATAACCCAACAAGCCCCCGCTATGGTGGTATTTCCTTAGGTTCAT
        ATGGCGGCTTTTGTTTCCATTTGATCTTCACAGCAATTCTCTACAGGAATCTGGGCAGAT
        TTATTTCCTTTAGAGGAATTTCCAGGTCTTAAAATCTATAGGGGGCAACTATCAAAACTT
        [C,T]
        ACCCAATGTTGCCCCCTACCCACACACAAAACCAGGCCCCCAGCCGATCAGAAAGCACTG
        CTGAGCTCCTGTCAGGGCCCACGCAGCTCGCTGTGAGACAGAGAGAGGGAACTCACATTT
        ATTGATCACCTACTGAGCATCCATCACTAGGCTAGGACCGTCACATTCCTTAACTTTTGA
        ATCCTTTCATGAGGTAGGCATTATTATTCTCCTTTTGTTTCACATAGCCATTAAAGAACA
        AAATTTGGGGCTGGGTGTGCTGACTCACACCTGTGATCTAGCACTTTAGGGGGCTGAGGC

13198   CTAACTATTCAGGAAGGTTAGGCGGGAGCACAACTTGGGTTCCAGGGTTTGAGGCTCCAG
        TGAGCTGATCTTGCCACTGCACTACAGCCTGAGCAACAGAGCAAGACCCTGTGACTCCAA
        AAACAAACAAACAAACACATTTTTGAACCCAAACAGATCTGACCCAAGATGCATGCTCTTA
        TAGATGCCACCTCCCTGTGTGCTGGGGCTTCTACTAAAAACACAGACAAGATCAGGCAAC
        CACAGTCAATCTAAGGGAAAGAGGAAAGTGTAACCAAAGCACAAATACATAAAATATTGC
```

FIGURE 3M

[-,A]
AAAAATGCTATTTAAAGAAAAAAAAAGAGAAGAGAGGCTCTGAGGTTGTACTAACAGAGAAT
GGCCTTGGCTAATCCAGGAAGACTTCCTGAAAGAGGTTGTTTTTTCCCCAGGTCTGCTTT
TGACATCTCTCTTTTCACAGTGCATCTGGGTAGTGAGCTTCCTCTCCTCCTTCTTCCTCA
GCCTGCCCTATGGTGTGGCAGTGGGTGTCGCCTTCTCCGTCCTGGTCGTGGTCTTCCAGA
CTCAGTTGTAAGTGATAGCTTCCGCCCTCCTAGGCCCACAGTCGGTTCCCTGGGCCAGCC

13257   GTGAGCTGATCTTGCCACTGCACTACAGCCTGAGCAACAGAGCAAGACCCTGTGACTCCA
AAAACAAACAAACAAACACATTTTGAACCCAAACAGATCTGACCCAAGATGCATGCTCTT
ATAGATGCCACCTCCCTGTGTGCTGGGGCTTCTACTAAAAACACAGACAAGATCAGGCAA
CCACAGTCAATCTAAGGGAAAGAGGAAAGTGTAACCAAAGCACAAATACATAAAATATTG
CAAAAATGCTATTTAAAGAAAAAAAAAGAGAAGAGAGGCTCTGAGGTTGTACTAACAGAGA
[A,G]
TGGCCTTGGCTAATCCAGGAAGACTTCCTGAAAGAGGTTGTTTTTTCCCCAGGTCTGCTT
TTGACATCTCTCTTTTCACAGTGCATCTGGGTAGTGAGCTTCCTCTCCTCCTTCTTCCTC
AGCCTGCCCTATGGTGTGGCAGTGGGTGTCGCCTTCTCCGTCCTGGTCGTGGTCTTCCAG
ACTCAGTTGTAAGTGATAGCTTCCGCCCTCCTAGGCCCACAGTCGGTTCCCTGGGCCAGC
CCGCAAAGGGCTTCCATGCCACGGCCTGGCTTAGTCCACTGTACCTTCCACCTCTGGGCC

14541   TCATGGACACTGACATTTATGTGAATCCCAAGACCTATAATAGGGTAGGTAATTCAAGCT
TATGACCTCCTTCTTTTGCTCTGCACCACCCCAAGAAGAGGTTGCTTTTTAAAGCCAATA
AAGACATTTCTGCAACTTGAGCTCAGTCTCCCTGTCACAGGCCCAGGATATCCAGGGGAT
TAAAATCATCACGTACTGCTCCCCTCTCTACTTTGCCAACTCAGAGATCTTCAGGCAAAA
GGTCATCGCCAAGGTAAGGCTCAGTCCCTGGCGACCAGAGGCTCTGGACAGAGAGTGGCC
[G,A]
GAAAATGGAAGCAGAAGGGCGGTGGGAGCTGAGAATAGGCCACTCCCATAGAGGGTGGAG
GTCAAGATTGCTGTTGGCTCTCTCCCTGCAGACAGGCATGGACCCCCAGAAAGTATTACT
AGCCAAGCAAAAATACCTCAAGAAGCAGGAGAAGCGGAGAATGAGGCCCACACAACAGAG
GAGGTCTCTATTCATGAAAACCAAGGTGAATGAAGGCCAGAAGCAGCCCCGTGCCCTGCT
CTCCTGCCCATTCTGATACTGCCCCCTGTTACTCATGGTACCCTGGGGGCCCCGCTTCCC

14545   GGACACTGACATTTATGTGAATCCCAAGACCTATAATAGGGTAGGTAATTCAAGCTTATG
ACCTCCTTCTTTTGCTCTGCACCACCCCAAGAAGAGGTTGCTTTTTAAAGCCAATAAAGA
CATTTCTGCAACTTGAGCTCAGTCTCCCTGTCACAGGCCCAGGATATCCAGGGGATTAAA
ATCATCACGTACTGCTCCCCTCTCTACTTTGCCAACTCAGAGATCTTCAGGCAAAAGGTC
ATCGCCAAGGTAAGGCTCAGTCCCTGGCGACCAGAGGCTCTGGACAGAGAGTGGCCGGAA
[A,G]
ATGGAAGCAGAAGGGCGGTGGGAGCTGAGAATAGGCCACTCCCATAGAGGGTGGAGGTCA
AGATTGCTGTTGGCTCTCTCCCTGCAGACAGGCATGGACCCCCAGAAAGTATTACTAGCC
AAGCAAAAATACCTCAAGAAGCAGGAGAAGCGGAGAATGAGGCCCACACAACAGAGGAGG
TCTCTATTCATGAAAACCAAGGTGAATGAAGGCCAGAAGCAGCCCCGTGCCCTGCTCTCC
TGCCCATTCTGATACTGCCCCCTGTTACTCATGGTACCCTGGGGGCCCCGCTTCCCACCC

15041   ACCAAGGTGAATGAAGGCCAGAAGCAGCCCCGTGCCCTGCTCTCCTGCCCATTCTGATAC
TGCCCCCTGTTACTCATGGTACCCTGGGGGCCCCGCTTCCCACCCTGACAGGCAAAGACA
GAAAGTCTCTGGGAACACTGCCTGGTGGCCGCTGGGCATTTTTCTTCTTTTTTTTCTTTT
TCTTTTTAGAGATGGAATTTTGCTCTTGTCACCCAGGCTTGAGTGCAATGGCGTTATCTT
GGCTCACTGCAACCTCCACCTCTGGGGTTCAAGCGATTCTCCTGCCTTAGCCTCCCAAGT
[C,A]
GCTGAGATTACAGGTGCCACCACACCCAGCTAATTTTTGTATTTTTAGTAGATATTGGGT
TTCACCATGTTGGCCAGGCTGGTGTCAAACTCCTGACCTCAGGTGATCCACCTACCTTAG
CCTTCCAAAGTGCTGGGATTACAAGCCTGAGCCACTGCGCCCAGCCTGGGCATTTTTCTT
CTTGGATGAGGTGCTACCATCTCCCAGGGAAGCCACTGAACCCCCAAGGCCCTTCTCCAT
TTTCTGGCTAAGATAGGACATGGCCCATGGACTTTTGAACAACCCAGAGGGGGAACAGCA

15053   GAAGGCCAGAAGCAGCCCCGTGCCCTGCTCTCCTGCCCATTCTGATACTGCCCCCTGTTA
CTCATGGTACCCTGGGGGCCCCGCTTCCCACCCTGACAGGCAAAGACAGAAAGTCTCTGG
GAACACTGCCTGGTGGCCGCTGGGCATTTTTCTTCTTTTTTTTTCTTTTTCTTTTTAGAGA
TGGAATTTTGCTCTTGTCACCCAGGCTTGAGTGCAATGGCGTTATCTTGGCTCACTGCAA
CCTCCACCTCTGGGGTTCAAGCGATTCTCCTGCCTTAGCCTCCCAAGTCGCTGAGATTAC
[A,C]
GGTGCCACCACACCCAGCTAATTTTTGTATTTTTAGTAGATATTGGGTTTCACCATGTTG
GCCAGGCTGGTGTCAAACTCCTGACCTCAGGTGATCCACCTACCTTAGCCTTCCAAAGTG
CTGGGATTACAAGCCTGAGCCACTGCGCCCAGCCTGGGCATTTTTCTTCTTGGATGAGGT
GCTACCATCTCCCAGGGAAGCCACTGAACCCCCAAGGCCCTTCTCCATTTTCTGGCTAAG
ATAGGACATGGCCCATGGACTTTTGAACAACCCAGAGGGGGAACAGCAGTGAATTTCCTG

15065   CAGCCCCGTGCCCTGCTCTCCTGCCCATTCTGATACTGCCCCCTGTTACTCATGGTACCC

FIGURE 3N

```
       TGGGGGCCCCGCTTCCCACCCTGACAGGCAAAGACAGAAAGTCTCTGGGAACACTGCCTG
       GTGGCCGCTGGGCATTTTTCTTCTTTTTTTTCTTTTTCTTTTTAGAGATGGAATTTTGCT
       CTTGTCACCCAGGCTTGAGTGCAATGGCGTTATCTTGGCTCACTGCAACCTCCACCTCTG
       GGGTTCAAGCGATTCTCCTGCCTTAGCCTCCCAAGTCGCTGAGATTACAGGTGCCACCAC
       [A,G]
       CCCAGCTAATTTTTTGTATTTTTAGTAGATATTGGGTTTCACCATGTTGGCCAGGCTGGTG
       TCAAACTCCTGACCTCAGGTGATCCACCTACCTTAGCCTTCCAAAGTGCTGGGATTACAA
       GCCTGAGCCACTGCGCCCAGCCTGGGCATTTTTCTTCTTGGATGAGGTGCTACCATCTCC
       CAGGGAAGCCACTGAACCCCCAAGGCCCTTCTCCATTTTCTGGCTAAGATAGGACATGGC
       CCATGGACTTTTGAACAACCCAGAGGGGGAACAGCAGTGAATTTCCTGGGGAACCCAGGC

15108  TGTTACTCATGGTACCCTGGGGGCCCCGCTTCCCACCCTGACAGGCAAAGACAGAAAGTC
       TCTGGGAACACTGCCTGGTGGCCGCTGGGCATTTTTCTTCTTTTTTTTCTTTTTCTTTTT
       AGAGATGGAATTTTGCTCTTGTCACCCAGGCTTGAGTGCAATGGCGTTATCTTGGCTCAC
       TGCAACCTCCACCTCTGGGGTTCAAGCGATTCTCCTGCCTTAGCCTCCCAAGTCGCTGAG
       ATTACAGGTGCCACCACACCCAGCTAATTTTTTGTATTTTTAGTAGATATTGGGTTTCACC
       [A,C]
       TGTTGGCCAGGCTGGTGTCAAACTCCTGACCTCAGGTGATCCACCTACCTTAGCCTTCCA
       AAGTGCTGGGATTACAAGCCTGAGCCACTGCGCCCAGCCTGGGCATTTTTCTTCTTGGAT
       GAGGTGCTACCATCTCCCAGGGAAGCCACTGAACCCCCAAGGCCCTTCTCCATTTTCTGG
       CTAAGATAGGACATGGCCCATGGACTTTTGAACAACCCAGAGGGGGAACAGCAGTGAATT
       TCCTGGGGAACCCAGGCAGCCCAGGGCTAGCAAGGCTGGGGTGGCCATGGCAGTAATCCT

16274  CTTCCAGACTGTCTCCCTGCAGGAGCTGCAGCAGGACTTTGAGAATGCGCCCCCCACCGA
       CCCCAACAACAACCAGACCCCGGCTAACGGCACCAGCGTGTCCTATATCACCTTCAGCCC
       TGACAGCTCCTCACCTGCCCAGAGTGAGCCACCAGCCTCCGCTGAGGCCCCCGGCGAGCC
       CAGTGACATGCTGGCCAGCGTCCCACCCTTCGTCACCTTCCACACCCTCATCCTGGACAT
       GAGTGGAGTCAGCTTCGTGGACTTGATGGGCATCAAGGCCCTGGCCAAGGTGAGGCCCTC
       [-,G]
       GGGACAGCAAGCACCACCCACTCCACCCCCTCCGCTCTGCTCTCCACATTCCCTTTCCTG
       GGAGCCCTCATTTCAGGAAGCTGAGGGAGGAAGCTCACTGGGGAGGCTAACAGCTCCTAG
       GAATCCCTCCTTTCCCCAGACGCCCACCAGGTTGAGACATTCTCCACAGAGCAGGCCCAGA
       CGGCCCATGACAATGAGTGGCGGGACAAGTCTACCAGAGTTTCAGGCCCCTGTGCTCCCA
       ACACCCCCAGCAGTGGCCATCCCAAGTCCCTCTCAGCCATCAGGAACCCACCCAGGTTCT

17424  AACATGGTGAAACCCCGCCTCTACTAAAAATACAAAAATTAGCCAGGTGTGGTGACGGGC
       CCCTGTAGTCCCAGCTACTCGGTAGGCTGAGGCAGAGAATTGCTTGAACCCAGGAGGCGG
       AGGTTGCAGTGAGCCAAGATCGCGCCACTGCACTCCAGCCTGGGCAACAGAGTGAAACTC
       CATCTCAAAAGAAAAAAAAAGAAAATATCTAGCCCCACAAGAAGGGGCCATGGTGACTTT
       AAGTGCCCGCCACGTTGGCAAAAGTCCATTTCCGCTCCACTTCCCAGAGAAACCGTCAGC
       [C,T]
       AACACTCCAGGGAGAAGTGGTGTGCTTTGCTGCTATTTTTGTCTTTGGCTGCTGGGCTCT
       CAGGGTTGCTTATTTGTTTGGCTTCCCCTCTGAAGTACGTTTTGTGAATCACTTTTGAGA
       CCCACTCAGAACATTCCTTTCCTTTTGCCTCCCTACCCCAACAACACTTCTAGCTGAGCT
       CCACCTATGGGAAGATCGGCGTGAAGGTCTTCTTGGTGAACATCCATGGTAAGAGAAAGA
       GGACATTTAGGGACTGAAAGACTGGCAAGGAGTGTGGGGTAGGAACAGGTTGGTGGGGTC

17627  AATATCTAGCCCCACAAGAAGGGGCCATGGTGACTTTAAGTGCCCGCCACGTTGGCAAAA
       GTCCATTTCCGCTCCACTTCCCAGAGAAACCGTCAGCCAACACTCCAGGGAGAAGTGGTG
       TGCTTTGCTGCTATTTTTGTCTTTGGCTGCTGGGCTCTCAGGGTTGCTTATTTGTTTGGC
       TTCCCCTCTGAAGTACGTTTTGTGAATCACTTTTGAGACCCACTCAGAACATTCCTTTCC
       TTTTGCCTCCCTACCCCAACAACACTTCTAGCTGAGCTCCACCTATGGGAAGATCGGCGT
       [G,A]
       AAGGTCTTCTTGGTGAACATCCATGGTAAGAGAAAGAGGACATTTAGGGACTGAAAGACT
       GGCAAGGAGTGTGGGGTAGGAACAGGTTGGTGGGGTCTGAATAGTGAGGAGGTTGGAAAC
       GAGAGCACCCAGCTATCCCCCACAAGCTGCTGCCTGCTCATAAAAGCTTCAGGTACAAGT
       CCAAAGAGACTGGTCAGATTGCATAAACATCCTAGGGGCCTTAGTGACAGAGTGGGGGTG
       AGGAGGTCATGGAGTTACAGAAGGACAGCTAGGATTCTAATCTACCCCATAACTAATTTG

18427  GGGTGCATATACACAGCCTCAAGGACGTGGCCACAGGGCAGCAGACATTTACATGACTAG
       CATGTACGCAAAGTGCAGAGATGTGGGAGCAAGTGCACACAGACACACAGGAGAATGTGA
       AGGGGCACATACACACACCCAGCTCCCTGCACTGGGTCAGACCCCCTCCAGCAGGGCT
       GCAGTTCCCAAGCTCCGCATGGCCACGTTCGGGGAGAGAATCTGCAGTGGCAATGACCTG
       CTATGATATGTTCTGGAGTTAGAAGCAGTGGATTCTCCCCAACCTCACTGGACACCCCCT
       [T,C]
       AGGAAAACCATCTCTAGGATTAAGAGTAATCCACACAAACTTCCAATGCCACACATTGGAA
       GTTGCTGGAAAGGTCTGGGAAAACAAGAGGAAGGATGGGTCCTTGGGGGATAGAACTGGC
       AGCGGCCTCTTCAAGGATGGCTTAGGCTTTTCCACTCGAATCACCACAAAGTACTGACTC
```

FIGURE 30

CCTAAATCAAACTGCTTCCTTCTGCTCTGGGTTGAAACTTCAGCATCCTCAAGTTCATGT
TGCCCTCTGCCGTCCAGAACTGATATTGCACTGCCAATGCCATGGCCCTCAGATACAGCA

18813 AGAGGAAGGATGGGTCCTTGGGGGATAGAACTGGCAGCGGCCTCTTCAAGGATGGCTTAG .
GCTTTTCCACTCGAATCACCACAAAGTACTGACTCCCTAAATCAAACTGCTTCCTTCTGC
TCTGGGTTGAAACTTCAGCATCCTCAAGTTCATGTTGCCCTCTGCCGTCCAGAACTGATA
TTGCACTGCCAATGCCATGGCCCTCAGATACAGCAAGAGCTGGGACCTCAGGCCTCTCCC
ATCCCTGCTCTGGTCTCACTATCTTCCCCACCCCCAGCTCCAATCCACAATGGCTGTTAT
[C,G]
TTTCTGAAGGTGATCTTTTCTCCTTCTAGCCCAGGTGTACAATGACATTAGCCATGGAGG
CGTCTTTGAGGATGGGAGTCTAGAATGCAAGCACGTCTTTCCCAGCATACATGACGCAGT
CCTCTTTGCCCAGGCAAATGCTAGAGACGTGACCCCAGGACACAACTTCCAAGGGGTAAG
GTTCTTGCACCTGGGGAATCCTAGGCTCCAAGGCACTGAAATAGCAGGACCAAGAGGCAT
TATTAGAAAGAACACAGGAGAAGGTTTAAGTTCCAATATCAAGTCTGCCATTTCAGTTTT

19035 GGACCTCAGGCCTCTCCCATCCCTGCTCTGGTCTCACTATCTTCCCCACCCCCAGCTCCA
ATCCACAATGGCTGTTATCTTTCTGAAGGTGATCTTTTCTCCTTCTAGCCCAGGTGTACA
ATGACATTAGCCATGGAGGCGTCTTTGAGGATGGGAGTCTAGAATGCAAGCACGTCTTTC
CCAGCATACATGACGCAGTCCTCTTTGCCCAGGCAAATGCTAGAGACGTGACCCCAGGAC
ACAACTTCCAAGGGGTAAGGTTCTTGCACCTGGGGAATCCTAGGCTCCAAGGCACTGAAA
[T,C]
AGCAGGACCAAGAGGCATTATTAGAAAGAACACAGGAGAAGGTTTAAGTTCCAATATCAA
GTCTGCCATTTCAGTTTTCTGAATCTGTTTCCTTATCTATAGAATGAGCACCATCAACTA
ACATTACCTACCTCTCTGCATTTTTCTTTTATTTTGTTTTAGGGTTAAATGATAATTACA
TCTTTTGTGTCACTTGAAAGCACTTTGTGTATTGTAAAAATTCTTTATCAATATAAGTTT
TCTGGTTGCACAAACACCCAAAGCATAGTAGAGCAGGCCCACTCTGCTGGCATCGTTCCC

19182 AGGATGGGAGTCTAGAATGCAAGCACGTCTTTCCCAGCATACATGACGCAGTCCTCTTTG
CCCAGGCAAATGCTAGAGACGTGACCCCAGGACACAACTTCCAAGGGGTAAGGTTCTTGC
ACCTGGGGAATCCTAGGCTCCAAGGCACTGAAATAGCAGGACCAAGAGGCATTATTAGAA
AGAACACAGGAGAAGGTTTAAGTTCCAATATCAAGTCTGCCATTTCAGTTTTCTGAATCT
GTTTCCTTATCTATAGAATGAGCACCATCAACTAACATTACCTACCTCTCTGCATTTTTC
[T,C]
TTTATTTTGTTTTAGGGTTAAATGATAATTACATCTTTTGTGTCACTTGAAAGCACTTTG
TGTATTGTAAAAATTCTTTATCAATATAAGTTTTCTGGTTGCACAAACACCCAAAGCATA
GTAGAGCAGGCCCACTCTGCTGGCATCGTTCCCTGCCTCCTCCTCATCTCTTTCTAAAGG
GGGCTTTCGGGAAGGGAGGGGAGGGGAGTAAGCCTACCCATTTTAACTTACCGGAGCTTA
GAGATTTCAGGCTGGTGAGGGATAAAGAGATTGGGTCTGAGTTTTGTCTCAGCTTTTTGA

19508 TAATTACATCTTTTGTGTCACTTGAAAGCACTTTGTGTATTGTAAAAATTCTTTATCAAT
ATAAGTTTTCTGGTTGCACAAACACCCAAAGCATAGTAGAGCAGGCCCACTCTGCTGGCA
TCGTTCCCTGCCTCCTCCTCATCTCTTTCTAAAGGGGGCTTTCGGGAAGGGAGGGGAGGG
GAGTAAGCCTACCCATTTTAACTTACCGGAGCTTAGAGATTTCAGGCTGGTGAGGGATAA
AGAGATTGGGTCTGAGTTTTGTCTCAGCTTTTTGACATTTAATTTACTAGCTCAGTAAGT
[-,G,C]
ATACAAATGGGATACAAATAACACCATCTAAAACTCCAGAAGACTGGGGAGTCAGAAAAA
TCCTACCTCCTTGGGGTCCCTGCCCAGATCCCCAGTCATCTCTAGCCCTCAGGGTCCCCT
CCCAGCTCAGCTCCTGCCCTTGGCCTCCCAAGACTCTTGTTGTGCCCCAGCCCTGGGTAA
AAACCTCCCCTGCCCTCTGTGGGTCATAAGAAAGGCTTTTCTGGCCCTAGAGCAATGATT
TGCTCTTTGCCTTAAGAGACTGATGAAGGTGAAACCATCTGTTCTAAGTGCTGAAAGACT

19571 AGTTTTCTGGTTGCACAAACACCCAAAGCATAGTAGAGCAGGCCCACTCTGCTGGCATCG
TTCCCTGCCTCCTCCTCATCTCTTTCTAAAGGGGGCTTTCGGGAAGGGAGGGGAGGGGAG
TAAGCCTACCCATTTTAACTTACCGGAGCTTAGAGATTTCAGGCTGGTGAGGGATAAAGA
GATTGGGTCTGAGTTTTGTCTCAGCTTTTTGACATTTAATTTACTAGCTCAGTAAGTCAT
ACAAATGGGATACAAATAACACCATCTAAAACTCCAGAAGACTGGGGAGTCAGAAAAATC
[T,G,C]
TACCTCCTTGGGGTCCCTGCCCAGATCCCCAGTCATCTCTAGCCCTCAGGGTCCCCTCCC
AGCTCAGCTCCTGCCCTTGGCCTCCCAAGACTCTTGTTGTGCCCCAGCCCTGGGTAAAAA
CCTCCCCTGCCCTCTGTGGGTCATAAGAAAGGCTTTTCTGGCCCTAGAGCAATGATTTGC
TCTTTGCCTTAAGAGACTGATGAAGGTGAAACCATCTGTTCTAAGTGCTGAAAGACTGCC
CAGGAACACACAGGGCGCTGGCTCCTGCCCTCCATGCCTAGAGGGAAACCCTGGGGAAAC

20147 GCCTAGAGGGAAACCCTGGGGAAACAACGGGCTTTCCTGCTTCGTGAAATTTGTCCGCAG
AGCAAAGAGGGAGATTCTGGAGGAAGCTGCATTAGTTGTTAGTGCCCTAATCATGTTCAG
CTACTCTAGTTGGTATGTATACTTGATTAGTCATAGCACTTATAAATAATTTATATTTTA
TATAATATATACTTACATATTATAGACCATTCACAGATACAAATCACACACATAAACACA
CACCTTTTTCAACAGCATTGTGAGGGACAAAGCAGGCAAAGTGAGGCTGGTTATCAGACTT

FIGURE 3P

[T,G]
AACAGATTAGAAAATATATTCCCAGGAGGACAGGAATTCCCCAAGGTCAGGCAGCTAGCC
AATAGTTTTTCTAAGCTGAGTAAAACCTTCCCTGCCTCTAACGGCCCACAAAGGAGGGAA
GACCGCGATACACACCTGTCTGGTATAAGGGGGAAGACCACAGCCGTGCTGTTTTTGTGA
GGCAGGTAAGGGAAGGGGCAAGAGGATAAGTCATGTGTCAGGAAGCAGCGTCCAACCAGA
GCCGGCCACCTGTCCCTTTTCCTGCCACCATGCACCAACTTTGCTGTTCAGTCACTGAAG

20180   TTCCTGCTTCGTGAAATTTGTCCGCAGAGCAAAGAGGGAGATTCTGGAGGAAGCTGCATT
AGTTGTTAGTGCCCTAATCATGTTCAGCTACTCTAGTTGGTATGTATACTTGATTAGTCA
TAGCACTTATAAATAATTTATATTTTATATAATATATACTTACATATTATAGACCATTCA
CAGATACAAATCACACACACATAAACACACACCTTTTCAACAGCATTGTGAGGGACAAAGCA
GGCAAAGTGAGGCTGGTTATCAGACTTTAACAGATTAGAAAATATATTCCCAGGAGGACA
[G,A]
GAATTCCCCAAGGTCAGGCAGCTAGCCAATAGTTTTTCTAAGCTGAGTAAAACCTTCCCT
GCCTCTAACGGCCCACAAAGGAGGGAAGACCGCGATACACACCTGTCTGGTATAAGGGGG
AAGACCACAGCCGTGCTGTTTTTGTGAGGCAGGTAAGGGAAGGGGCAAGAGGATAAGTCA
TGTGTCAGGAAGCAGCGTCCAACCAGAGCCGGCCACCTGTCCCTTTTCCTGCCACCATGC
ACCAACTTTGCTGTTCAGTCACTGAAGCTCATTCTGCACTGGCTTCCTCCCTTCCAGGCT

20584   TGTCTGGTATAAGGGGGAAGACCACAGCCGTGCTGTTTTTGTGAGGCAGGTAAGGGAAGG
GGCAAGAGGATAAGTCATGTGTCAGGAAGCAGCGTCCAACCAGAGCCGGCCACCTGTCCC
TTTTCCTGCCACCATGCACCAACTTTGCTGTTCAGTCACTGAAGCTCATTCTGCACTGGC
TTCCTCCCTTCCAGGCTCCAGGGGATGCTGAGCTCTCCTTGTACGACTCAGAGGAGGACA
TTCGCAGCTACTGGGACTTAGAGCAGGTGAGCTGAGGGAAGGGGCTGTGAGGGTGGGAGC
[A,T]
GGGCGAAGAGGGGAAGGATGGGGTCGCTGTCAAATACAAGGCGTTCACTCAGCTGTCTCA
CCTCCAGCCCAGAGCAGTCACATTCAAGGCCACAAAGATTTGTGGTCATCTTTGTTTTTT
TTCTTTTCCTTTTCTTTTTTTTTTTTTTTTTAATTTGAGACAAAGTCTCACTCTATCACCC
AGACTGGAATGCAGTGGCATGATCTCAGCTCACTGCAACCTCTGCCTCCCGGGTTCCAGA
GGTTCTCCTGCCTCAGCCTCCCGAGTAGCTGGGACTTCAGGCCTGCGCCCAGCTAATTTT

20717   ATGCACCAACTTTGCTGTTCAGTCACTGAAGCTCATTCTGCACTGGCTTCCTCCCTTCCA
GGCTCCAGGGGATGCTGAGCTCTCCTTGTACGACTCAGAGGAGGACATTCGCAGCTACTG
GGACTTAGAGCAGGTGAGCTGAGGGAAGGGGCTGTGAGGGTGGGAGCAGGGCGAAGAGGG
GAAGGATGGGGTCGCTGTCAAATACAAGGCGTTCACTCAGCTGTCTCACCTCCAGCCCAG
AGCAGTCACATTCAAGGCCACAAAGATTTGTGGTCATCTTTGTTTTTTTTTCTTTTCCTTT
[T,C]
CTTTTTTTTTTTTTTTTTTAATTTGAGACAAAGTCTCACTCTATCACCCAGACTGGAATGCA
GTGGCATGATCTCAGCTCACTGCAACCTCTGCCTCCCGGGTTCCAGAGGTTCTCCTGCCT
CAGCCTCCCGAGTAGCTGGGACTTCAGGCCTGCGCCCAGCTAATTTTTGTATTTTTAGTA
GAGACAGCTTTTCACCATGTTGGCTGGGCTGGTCTCGAACTTCCGATCTCAAGCAATCTG
CCTGCCTCGGTCTCCTAAGTGCCTGGATTACAGGCATAAGCCACGATGCCTGGCCTTTGT

20894   GGGGAAGGATGGGGTCGCTGTCAAATACAAGGCGTTCACTCAGCTGTCTCACCTCCAGCC
CAGAGCAGTCACATTCAAGGCCACAAAGATTTGTGGTCATCTTTGTTTTTTTTTTCTTTTCC
TTTTCTTTTTTTTTTTTTTTTTTAATTTGAGACAAAGTCTCACTCTATCACCCAGACTGGAA
TGCAGTGGCATGATCTCAGCTCACTGCAACCTCTGCCTCCCGGGTTCCAGAGGTTCTCCT
GCCTCAGCCTCCCGAGTAGCTGGGACTTCAGGCCTGCGCCCAGCTAATTTTTGTATTTTT
[A,G]
GTAGAGACAGCTTTTCACCATGTTGGCTGGGCTGGTCTCGAACTTCCGATCTCAAGCAAT
CTGCCTGCCTCGGTCTCCTAAGTGCCTGGATTACAGGCATAAGCCACGATGCCTGGCCTT
TGTTTTTCATTCTTCTCACTCCCTGAAAGGCATCGTGGGGAGAGGGTGAGTCACTGGACCA
AGTCCTAGAGAACCAGTATCTATTCTTATTCTCCAACACATCACCCACGTGACCCTGAGC
AAGCCACATACACCCTGGGCCCTAGTTTTTATCATCTGTGAAATTAGGGGAAACATAGGT

21787   GGGTGCAATGGTTCACACCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGCGGACCAC
CTGAGGTCAGGAGTTTGAGACCAGCCTGGCCAACATGGCGAAACCCCGTCTCTACTAAAA
GCACAAAAATTAGCCAGGCGTAGTGGTGCATGCCTGTAGTCCCAGCTACTCGGAAGTCTG
AGGCATGAGAATCACTTGAACCTGGGAGGCAGATGTTGCAGTGAGCCGAGATCGTGCCAC
TGCACTCCAGCTTGGGTGACAGAGCTAGACTGTCTCAAAAACAAACAAACAAAACAAAAC
[-,A,C]
TAAAAGATATGTGGATATGAGGGATCACCATCCCCATAGGGCCCCTGGATTAACACCACC
CCACCAATGCCCTGAATTAAAAGAAACCAGATGACTAGGTTTGGAGAAATCTGGCTTTGG
GTCTATGAGAAGTAGTGTCTCTCTTTGTGCCTCTTCCCATTCTTTTTTGACATTGAGCTCC
ATGGTGCTCTGAATCCGTCTCTCACAGTGCTGATGGCAGGTGGGACAGATTAGAAAATAG
AGCTGGAGCCACAGAGATTTGGCAGACTGATTTCGGTGCCCTCTTGGAATCTCCAGCACA

22264   CTCCATGGTGCTCTGAATCCGTCTCTCACAGTGCTGATGGCAGGTGGGACAGATTAGAAA

FIGURE 3Q

ATAGAGCTGGAGCCACAGAGATTTGGCAGACTGATTTCGGTGCCCTCTTGGAATCTCCAG
CACATTCCAAAAAGCCTGGATAGGACCAAAATAGCTTATCAACGTGAGAAAGGACTTCAG
AGCTTGTCTACTGCCAACCCTCATTTTACCCAATGAGGAAAGTGAAGCTATTAGGGGGCG
AGGGACACGTGGAAGGTCACACAGCACACAGGAGGTGATTCACATGTAGATTTCAGCACC
[T,C]
GCTCCTGCCACGCTGGACTGGTTCACCTCCTAGGCTGACCCTGCCTCTCCCCTGTTCACA
CACACTCTCGCACACACACACACACACACACACACACACAGGTGCTTTGTTCTGGCCAGG
GGTTCCTAGGGTCACCTCTTGGTTGCAGCCACTGTGACCCCAACTGGTCTAACCTCTCTC
TTCCCCTCCCACTTCCTTCCTGTGGTTCCTGCAGGAGATGTTCGGGAGCATGTTTCACGC
AGAGACCCTGACCGCCCTGTGAGGGCTCAGCCAGTCCTCATGCTGCCTACAGAGTGCCTG

22338   ACAGAGATTTGGCAGACTGATTTCGGTGCCCTCTTGGAATCTCCAGCACATTCCAAAAAG
        CCTGGATAGGACCAAAATAGCTTATCAACGTGAGAAAGGACTTCAGAGCTTGTCTACTGC
        CAACCCTCATTTTACCCAATGAGGAAAGTGAAGCTATTAGGGGGCGAGGGACACGTGGAA
        GGTCACACAGCACACAGGAGGTGATTCACATGTAGATTTCAGCACCTGCTCCTGCCACGC
        TGGACTGGTTCACCTCCTAGGCTGACCCTGCCTCTCCCCTGTTCACACACACTCTCGCAC
        [-,C,A]
        CACACACACACACACACACACACACACAGGTGCTTTGTTCTGGCCAGGGGGTTCCTAGGGTCA
        CCTCTTGGTTGCAGCCACTGTGACCCCAACTGGTCTAACCTCTCTCTTCCCCTCCCACTT
        CCTTCCTGTGGTTCCTGCAGGAGATGTTCGGGAGCATGTTTCACGCAGAGACCCTGACCG
        CCCTGTGAGGGCTCAGCCAGTCCTCATGCTGCCTACAGAGTGCCTGGCACTTGGGACTTC
        CATAAAGGATGAGCCTGGGGTCACAGGGGGTGTCGGGCGGAGGAAAGTGCATCCCCCAGA

23363   CAGGGACCATGTGCTCTCCACACCCAGGAGTCTAGGCCTTGGTAACTATGCGCCCCCCGT
        CCATCATCCCCAAGGCTGCCCAAACCACCACTGCTGTCAGCAAGCACATCAGACTCTAGC
        CTGGACAGTGGCCAGGACCGTCGAGACCACCAGAGCTACCTCCCCGGGGACAGCCCACTA
        AGGTTCTGCCTCAGCCTCCTGAAACATCACTGCCCTCAGAGGCTGCTCCCTTCCCCTGGA
        GGCTGGCTAGAAACCCCAAAGAGGGGGATGGGTAGCTGGCAGAATCATCTGGCATCCTAG
        [T,C]
        AATAGATACCAGTTATTCTGCACAAAACTTTTGGGAATTCCTCTTTGCACCCAGAGACTC
        AGAGGGGAAGAGGGTGCTAGTACCAACACAGGGAAAACGGATGGGACCTGGGCCCAGACA
        GTCCCCCTTGACCCCAGGGCCCATCAGGGAAATGCCTCCCTTTGGTAAATCTGCCTTATC
        CTTCTTTACCTGGCAAAGAGCCAATCATGTTAACTCTTCCTTATCAGCCTGTGGCCCAGA
        GACACAATGGGGTCCTTCTGTAGGCAAAGGTGGAAGTCCTCCAGGGATCCGCTACATCCC

23688   AAACTTTTGGGAATTCCTCTTTGCACCCAGAGACTCAGAGGGGAAGAGGGTGCTAGTACC
        AACACAGGGAAAACGGATGGGACCTGGGCCCAGACAGTCCCCCTTGACCCCAGGGCCCAT
        CAGGGAAATGCCTCCCTTTGGTAAATCTGCCTTATCCTTCTTTACCTGGCAAAGAGCCAA
        TCATGTTAACTCTTCCTTATCAGCCTGTGGCCCAGAGACACAATGGGGTCCTTCTGTAGG
        CAAAGGTGGAAGTCCTCCAGGGATCCGCTACATCCCCTAACTGCATGCAGATGTGGAAAG
        [G,A]
        GGCTGATCCAGATTGGGTCTTCCTGCACAGGAAGACTCTTTAACACCCTTAGGACCTCAG
        GCCATCTTCTCCTATGAAGATGAAAATAGGGGTTAAGTTTTCCATATGTACAAGGAGGTA
        TTGAGAGGAACCCTACTGTTGACTTGAAAATAAATAGGTTCCATGTGTAAGTGTTTTGTA
        AAATTTCAGTGGAAATGCACAGAAAATCTTCTGGCCTCTCATCACTGCTTTTCTCAAGCT
        TCTTCAGCTTAACAACCCCTTCCCTAACAGGTTGGGCTGGCCCAGCCTAGGAAAACATCC

24210   TCACTGCTTTTCTCAAGCTTCTTCAGCTTAACAACCCCTTCCCTAACAGGTTGGGCTGGC
        CCAGCCTAGGAAAACATCCCCATTTCTAACTTCAGCCAGACCTGCGTTGTGTGTCTGTGT
        GTTGAGTGAGCTGGTCAGCTAACAAGTCTTCTTAGAGTTAAAGGAGGGGGTGCTGGCCAA
        GAGCCAACACATTCTTGGCCCAGGAGCATTGCTTTTCTGTGAATTCATTATGCCATCTGG
        CTGCCAATGGAACTCAAAACTTGGAAGGCGAAGGACAATGTTATCTGGGATTCACCGTGC
        [A,C]
        CAGCACCCGAAGTGCCAAATTCCAGGAGGACAAGAGCCTTAGCCAATGACAACTCACTCT
        CCCCTACTCCACCTCCTTCCAAGTCCAGCTCAGGCCCAGGAGGTGGGAGAAGGTCACAGA
        GCCTCAGGAATTTCCAAGTCAGAGTCCCCTTTGAACCAAGTATCTAGATCCCCTGAGGAC
        TTGATGAAGTGATCCTTAACCCCCAAGTAATCATTAACCCCCAGACCAGCCTCAGAACTG
        AAGGAGATTGTTGACCCAGTGACCTGGAGTTGAGGCTCAGGGAGAGATCTGCCACATGTC

FIGURE 3R